ESSAYS IN FINANCIAL FRAGILITY

By

YULIYAN MITKOV

A Dissertation Submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Economics

Written under the direction of

Todd Keister

And approved by

____

____

____

____

New Brunswick, New Jersey

October 2017

ABSTRACT OF THE DISSERTATION

Essays in Financial Fragility

by YULIYAN MITKOV

Dissertation Director:

Professor Todd Keister

This dissertation is composed of three separate, but closely related, essays on financial instability. Chapter 1 offers new insights into the fragility-enhancing economic mechanisms at work during the Financial Crisis of 2007-08. Chapter 2 reexamines the effectiveness of recent regulatory measures aiming to mitigate future episodes of financial turmoil. Chapter 3 proposes a novel approach to an old problem in the literature on financial instability, namely how to derive sharper predictions in models with multiple equilibria.

In Chapter 1, I explore how the distribution of wealth across households influences the government's response to a banking crisis and the fragility of the financial system. In particular, I analyze a version of the Diamond and Dybvig (1983) model of financial intermediation where households have heterogeneous endowments and a government collects taxes and uses the proceeds to finance the provision of a public good. In addition, if there is a financial panic, the government can use some tax revenue to bail out banks experiencing a run. I show that when the wealth distribution is unequal, the government's bailout policy during a systemic crisis will be shaped in part by distributional concerns. In particular, government guarantees of deposits will tend to be credible for relatively poor investors, but may not be credible for wealthier investors. As a result, wealthier investors will have a stronger incentive to panic and,

in equilibrium, the institutions in which they invest are more likely to experience a run and receive a bailout. Thus bailouts, when they occur, will tend to benefit relatively wealthy investors at the expense of the general public. Notice that this result obtains naturally in my setting, without any appeal to political frictions or other factors that would give the wealthy undue influence over government policy. Rising inequality can strengthen this pattern. In particular, one of the effects of higher inequality is to make the panic-and-bailout cycle for the wealthy investors easier to obtain in equilibrium. In some cases, more progressive taxation reduces financial fragility and can even raise equilibrium welfare for all agents.

In Chapter 2, which is joint work with Todd Keister, we study the interaction between a government's bailout policy during a banking crisis and individual banks' willingness to impose losses on (or "bail in") their investors. Our interest in this topic is motivated by the fact that, in recent years, policy makers in several jurisdictions have drafted rules requiring financial institutions to impose losses on their investors in any future crisis. These rules aim both to protect taxpayers in the event of a future crisis and to change the incentives of banks and investors in a way that makes such a crisis less likely. While the specific requirements vary, and are often yet to be finalized, in many cases the bail-in will be triggered by an announcement or action taken by the institution facing losses. This fact raises the question of what incentives banks will face when deciding whether and when to bail in their investors. Banks in our model hold risky assets and are free to write complete, state-contingent contracts with investors. In the constrained efficient allocation, banks experiencing a loss immediately cut payments to withdrawing investors. In a competitive equilibrium, however, these banks often delay cutting payments in anticipation of being bailed out. In some cases, the costs associated with this delay are large enough that investors will choose to run on their bank, creating further distortions and deepening the crisis. We discuss the implications of the model for banking regulation and optimal policy design.

In Chapter 3, I investigate a new approach to endogenizing the probability of a self-fulfilling outcome in games of coordination. Specifically, a number of important economic phenomena such as currency attacks, bank runs and sovereign defaults can be understood as collective action problems where the players can end up coordinating on one of two different outcomes with markedly different consequences. This multiplicity of possible equilibrium outcomes presents a theoretical challenge since it renders the model predictions and its comparative statics relatively ambiguous. One approach to deriving sharper predictions in collective action problems is the global games framework initially proposed by Carlson and Van Damn (1993) and further developed by Frankel, Morris, and Pauzner (2000). The private sunspot approach is an alternative way of endogenizing the probability of a self-fulfilling event. The purpose of Chapter 3 is to illustrate the logic of the private sunspot approach through a simple example referred to as the Bandit Game. In particular, I analyze a coordination game where two bandits receive an idiosyncratic signal of the realization of a random variable and want to coordinate on attacking a village in order to seize whatever it had produced. By being unrelated to the fundamentals of the environment, this random variable adds uncertainty to the model that is purely extrinsic (i.e. a sunspot). I refer to the bandits' idiosyncratic signals of this random variable as private sunspots (as opposed to public sunspots, which are perfectly observed) and study equilibria where the strategies of the bandits are conditioned on their private sunspot signals. In other words, the private sunspot generalizes the public sunspot approach by introducing strategic uncertainty in the bandits' actions. I show that under certain condition, the private sunspot equilibrium involving an attack on the village will be unique, with the probability of an attack pinned down by the features of the environment.

Acknowledgements

First and foremost, I want to thank my advisor Professor Todd Keister. The work in this thesis would certainly not have been possible without his unending mentoring, encouragement and support during the years of my doctoral studies. I am also deeply grateful to the remaining members of my dissertation committee: Professor Roberto Chang and Professor Tomas Sjostrom, each of whom contributed to my work through constant encouragement in addition to numerous invaluable suggestions which had greatly improved the quality of this work. In addition, I am indebted to the outside member of my committee – Professor Huberto Ennis from the Federal Reserve Bank of Richmond for providing valuable feedback on this work.

During my time at Rutgers I have greatly benefited from interacting with a number of the faculty members, both in discussing my research ideas with them and in attending their lectures. In addition to the members of my dissertation committee, this includes Professor Richard McLean, Professor Eugene White, Professor Hugh Rockoff and Professor John Landon-Lane.

I want to express my gratitude to a big number of graduate students here at Rutgers: Walter Bazan, Wukuang Cun, Antonio Cusato, Hassan Nosratabadi, Vladimir Kolchin, Geoffrey Clarke and Maria Pagliari for numerous conversations over a wide range of topics. They have certainly made the last five years of my life more enjoyable. In addition, I want to express my gratitude to the "macro-finance people" Yang Lee, Hyeon Ok Lee, Ryuichiro Izumi and Yan Lee for providing such an intellectually stimulating environment during our numerous working group meetings.

Finally, I want to express my deepest gratitude to my family. They have always encouraged me to pursue my goals, even if this entails that I will be away for a long period of time.

Contents

List of Figures

List of Tables

Chapter 1

Inequality and Financial Fragility

1.1 Introduction

In most of the theoretical literature on banking panics, both deposit insurance and panics are all-or-nothing affairs, in the sense that all deposits are treated equally and a panic affects all banks the same way. However, the financial crises observed in reality are more complicated. Deposit insurance typically covers some types of deposits (particularly smaller retail deposits), but not others. Panics are often restricted to certain types of institutions or arrangements (money market mutual funds in the United States in 2008) while others remain effectively insured by the government (commercial banks in the United States in 2008). Even within a single institution, some depositors may be forced to accept a haircut, whereas others are protected in full by the government (Cyprus in 2008). In addition, the written rules of the deposit insurance program might be abandoned in a systemic financial collapse, so that a banking crisis transforms the government guarantees from a legal to political commitment (Sibert, 2013). These issues have led some observers to question the view that government deposit insurance can solve the problem of banking panics (Cooper and Kempf, 2015).[1]

In this paper, I study a version of the Diamond and Dybvig (1983) model of financial intermediation where the distribution of investors' wealth influences the government's response during financial crisis, which in turn determines banks' sus-

[1] A case in point is the Icesave dispute, taking place after the collapse in 2008 of the Icelandic bank Landsbanki. The Court of Justice of the European Free Trade Association States (EFTA) ruled that the Icelandic government was not under the legal obligation to adhere to its original promise to insure Dutch and UK depositors, since doing so would have undermined the stability of the Icelandic banking system.

ceptibility to financial panics. Introducing wealth inequality allows me to capture the heterogeneity across types of investors and financial intermediaries that appears to be an important factor behind the government's emergency response to a crisis. As in Keister (2015), fiscal policy is introduced via a government that collects taxes and uses the proceeds to finance the provision of a public good or to bail out banks during a panic. Importantly, the government cannot pre-commit to the details of the bailout plan before the crisis, but instead chooses the bailout policy after the crisis is already underway. I restrict attention to self-fulfilling financial panics since doing so allows me to highlight the main implications of the model in a clear and concise way.

I begin by showing that the ex-post optimal bailout policy generates *endogenous caps to deposit insurance* and these caps will be a decreasing function of investors' wealth levels.[2] Thus, when wealth is unequally distributed, the bailout policy of the government will impose larger haircuts on the wealthy investors. At the same time, expecting larger losses in a financial crisis, the wealthy will be more prone to panic and run on the banks - an event which leads the government to bail them out - thus creating a self-fulfilling panic-and-bailout cycle. Importantly, the model predicts that the wealthy investors will be *endogenously* more likely to receive bailouts in equilibrium, even though they have no particular power or inside connections. In fact, throughout the paper, I assume that the government is both utilitarian and benevolent. An equilibrium where the wealthy investors panic and receive a bailout is possible because the government will not find it ex-post optimal to provide a rescue package that is sufficient to prevent them from panicking in the first place.

The second contribution of the paper is to shed new light on the link between widening inequality and financial fragility. One of the effects of higher inequality is to make the panic-and-bailout cycle for the wealthy investors easier to obtain in

[2] In this paper, deposit insurance is broadly interpreted to include different forms of ex-post bailouts to banks, regardless of whether the actual transfer of public funds to the financial sector stems from an explicit or an implicit government guarantee.

equilibrium. At the same time, higher inequality makes investors with lower levels of wealth less prone to panic. This result might appear surprising at first, but the mechanism leading to this outcome within the model is relatively straightforward. A more unequal wealth distribution strengthens the utilitarian concerns of the government and therefore leads to a bailout intervention imposing relatively larger haircuts on the wealthy and lower haircuts on the poor. Thus, when inequality is sufficiently high, the utilitarian incentives of the government will prevent the poor from panicking and, at the same time, will lead to an equilibrium characterized by a panic and a subsequent bailouts for the wealthy investors.

Widening inequality could also have a negative fiscal capacity effect, since a government with limited re-distributive power will respond to rising inequality by collecting less tax revenue. The reason is that keeping the same tax revenue stream would place a large burden on the increased fraction of relatively poor agents who have high marginal utility of consumption and therefore high marginal cost of paying an additional dollar in taxes. At the same time, a fall in tax revenue has a negative effect on financial stability since the government's tax revenue is related to the stability of the financial system. In particular, a government that collects less tax revenue has a lower capacity to rescue the banking sector, which in turn will make investors more prone to panic and run on the banks.

Finally, the analysis in the paper provides a novel justification for progressive taxation in times of rising inequality. In particular, if the promise of deposit insurance fails to be credible for wealthy investors, then they will be susceptible to financial panics. One way to restore credibility is to redistribute some of their wealth. This redistribution could prevent the self-fulfilling panic and bailout cycle for wealthy investors and therefore be desirable, provided that the efficiency loss from the increased progressivity is not prohibitively large. In fact, wealthy investors might also be willing to accept a more progressive tax code if this is the cost of making the government's guarantees credible.

Related literature. This paper is related to the recent work of Cooper and Kempf (2015), who study the government' ex post decision to provide deposit insurance in a version of the Diamond and Dybvig model where depositors have heterogeneous wealth levels. They show that ex-post, the government will abstain from providing deposit insurance if it entails high levels of undesirable redistribution. Copper and Kempf assume that the government must either provide deposit insurance to everyone or abstain from providing deposit insurance completely. In contrast, I allow for more flexible bailout interventions in which the government may choose to impose different haircuts on different investors, as often occurs during a systemic crisis. In addition, I analyze a model where agents incorporate the ex-ante probability of a bank run into their decisions (as in Cooper and Ross, 1998, Peck and Shell, 2003, and others). This approach allows me to analyze bailout interventions in equilibria where runs occur with positive probability. Cooper and Kempf (2015), on the other hand, restrict their analysis to financial arrangements that do not reflect the possibility of a bank run and focus on equilibria where the ex-post intervention prevents the run from taking place in equilibrium.

One of the main contributions of the paper is to examine the link between widening inequality and financial instability. Stiglitz (2009, 2012) and Fitoussi and Saraceno (2010, 2011) posit that widening income disparity during the years preceding the 2007-08 Financial Crisis depressed aggregate demand since it redistributed income from those with high propensity to consume to those with low propensity to consume (i.e. from low to high income individuals). The monetary policy response was to prop up aggregate demand by a prolonged period of low interest rates, which in turn set the stage for a credit expansion followed by the subsequent bust. According to Rajan (2010) increased inequality unleashes political pressures for more redistribution. This redistribution could take on forms - like subsidized lending for the poor – that introduce distortions and compromise the stability of the financial system. Acemoglu (2011) argues that rising inequality does not lead to redistributive policies benefiting the low and middle income voters, but instead to policies - like financial deregulation – which tend to disproportionately benefit the wealthy. This development arises because of political frictions, whereby the government is swayed by the preferences of the minority of high income voters, instead of the preferences of the majority of low and middle income voters. Kumhof and Rancière (2011) analyze a closed economy DGSE model where increased inequality pushes households in the lower part of the income distribution to become more indebted in order to maintain their standards of living, which in turn, makes the financial system less stable.

The model in this paper does not rely on political factors to derive a link between widening inequality and financial instability. Instead, a financial crisis is a form of coordination failure between investors. Whether such a coordination failure is possible in equilibrium for a given type of investors depends crucially on the way the bailout intervention of the government will treat different type of investors in case their panic. Changes in the distribution of wealth are therefore linked to financial fragility through their effect on the government's bailout policy.

One of the key assumptions in the paper is that the bailout policy is chosen in an ex-post efficient way after the advent of the crisis. The goal of this assumption is to capture the renegotiation of government guarantees that appears to play a major role in times of systemic banking failures. Ex-post efficient bailout policy has been analyzed by, among others, Chari and Kehoe (2010), Farhi and Tirole (2012), Nosal and Ordonez (2013), Ranciere and Tornell (2011) and Bianchi (2012) and Keister (2015).

The rest of the paper is organized as follows. The model is presented in section 1.2. Section 1.3 derives the properties of the panic equilibrium, while Section 1.4 analyzes the effect of increasing inequality on financial fragility. Section 1.5 augments the model to allow for a more general process by which investors choose where to place their deposits. Section 1.6 concludes. All proofs are collected in the appendix.

1.2 The Model

I extend a version of the Diamond and Dybvig model with fiscal policy and no commitment, as in Keister (2015), to include heterogeneous wealth levels among the investors.

1.2.1 The environment

Investors. There are three time periods $t = 0, 1, 2$ and a continuum of *investors* indexed by $i \in [0, 1]$. Each investor has preferences characterized by

$$u\left(c_1^i + \omega_i c_2^i\right) + v(g) = \frac{\left(c_1^i + \omega_i c_2^i\right)^{1-\gamma}}{1-\gamma} + \delta \frac{g^{1-\gamma}}{1-\gamma}, \quad \gamma > 1 \text{ and } \delta > 0$$

where c_t^i is the consumption of the private good in period $t = 1, 2$ and g is the level of the public good, which is provided in period 1. For each investor, the parameter ω_i can only take values in $\{0, 1\}$. If $\omega_i = 0$, investor i is *impatient* and values consumption only in period 1. If $\omega_i = 1$, investor i is *patient* and values consumption equally in period 1 and 2. All investors have the same probability π of being impatient and π is also the fraction of impatient investors in the population.

Endowments. At the beginning of period 0, the only difference between investors is their initial endowment e. In particular, I define a function E which maps each investor $i \in [0, 1]$ to his endowment $e \in [e_L, e_H]$:

$$E : [0, 1] \to [e_L, e_H]$$

Investors with endowment e will be called *type* e *investors*. The function E is common knowledge and the average endowment, $\int_{e_L}^{e_H} edG(e)$, is normalized to 1, where $G(e)$ denotes the fraction of investors whose endowment is less or equal to e.

Technology. Following Diamond and Dybvig (1983), there is a single, constant-return-to-scale technology, operated at a central location, which takes one unit of endowment in the initial period and transforms it into 1 unit of private consumption if liquidated in period 1. Every unit of endowment placed in the technology and not liquidated in the intermediate period yields $R > 1$ units of private consumption in the final period.

Sequential service. Investors' opportunity to contact the central location in order to withdraw arrives sequentially in a randomly determined order. Ex-ante, investors are equally likely to occupy any position in the order of opportunities to withdraw. At the beginning of period 1, each investor i learns whether he is patient or impatient and, in addition, his position $l(i) \in [0, 1]$ in the order of opportunities to withdraw. An investor with $l = 0$ knows that he is the first with an opportunity to withdraw, whereas an investor with $l = 1$ knows that he will be the last with an opportunity to withdraw. Each investor's order in the opportunities to withdraw is private information. When an investor's opportunity to withdraw arrives, he can either contact the central location and receive his payment or wait until the final period to withdraw. Investors are isolated from each other and those that withdraw in period 1 must consume immediately what is given to them and return to isolation. Wallace (1989, 1990) shows that this environment generates a *sequential service constraint* where the consumption of an investor depends only on the information available to the intermediation scheme at the time he withdraws.[3]

[3] This is different from the original approach of Diamond and Dybvig (1983), where after deciding to withdraw in period 1 each investor is randomly assigned a position in the withdrawing order. As shown in Green and Lin (2003), Andolfatto et. al. (2007) and Enis and Keister (2009b), investors' information about their position in this order plays an important role in the type of bank runs that can occur in equilibrium.

1.2.2 The decentralized economy

The constant-return-to-scale technology is operated by a continuum of banks. There is also a government who taxes investors' endowments at the beginning of period 0 and then uses the tax proceeds to provide a public good in period 1 and to make fiscal transfers (*bailouts*) to banks experiencing a run. Each investor chooses his withdrawing strategy as a part of a non-cooperative game between investors, the banks and the government. The possibility of financial panics is introduced via a sunspot state.

Banks. Banks perform an intermediation service by pooling investors' resources in the constant-return-to-scale technology with the goal of insuring them against idiosyncratic liquidity shocks. The banks observe investors' initial endowment and their choice to withdraw in the intermediate or in the final period. However, banks dot not observe whether a given investor is patient or impatient nor his position in the order of opportunities to withdraw and, therefore, payments cannot be made contingent on this information. In addition, banks cannot commit to a future plan of action. Instead, the payment to each investor is determined as a best response to the available information at the time of the withdrawal. The assumption of no-commitment is crucial for the results to follow. In contrast, if banks could pre-commit to their payment schedules, an equilibrium bank run will be prevented by suspending payments whenever more than a fraction π of the investors attempts to withdraw in period 1 (see Diamond and Dybvig (1983)).[4] Finally, banks behave competitively by taking economy-wide outcomes as independent of their actions and their objective is to maximize the expected utility of their investors at all times.

The government. The government is both *benevolent* and *utilitarian* and there-

[4]Ennis and Keister (2009a), on the other hand, show how lack of commitment can undermine a strict policy of suspension of convertibility and hence fail to prevent a run on the banks from taking place in equilibrium.

fore each investor is assigned the same weight in the social welfare function. The government collects taxes in period 0 by taxing investors' endowments. The tax rate τ must be the *same* for all investors and taxing is not possible after period 0. In period 1, the government allocates the tax proceeds between the provision of a public good and fiscal transfers to the financial sector (*bailouts*). The government is restricted to provide bailouts *only* to those banks experiencing a run. If τ denotes the tax proceeds and B the aggregate transfers to the banking sector, then the level of public good will be equal to $\tau - B$. The government cannot commit ex-ante to the details of the rescue plan and therefore will choose the bailout payments as a best response to the prevailing conditions at the time of the intervention.[5]

Financial panics. Following Cooper and Ross (1998), Peck and Shell (2003) and others, I allow investors to condition their withdrawal decisions on the realization of an extrinsic random variable s, which will be called the *sunspot state.* The sunspot state is unrelated to the fundamentals and has the interpretation of investor sentiment. The realization of s is observed by *all* investors at the beginning of period 1 and can on take two values, α and β, with respective probability $1-q$ and q. Henceforth, α is labeled the *good* state and β is labeled the *panic state.* Thus, I focus on the possibility that when $s=\beta$, it might become optimal for a patient investor to withdraw in period 1 (i.e. to *panic*) if he expects other patient investors to panic as well. A withdrawal strategy for investor i is a function y_i which specifies an action - either to withdraw in period 1 or in period 2 - for each possible combination of his preference type $\omega_i \in \{0,1\}$, the sunspot state $s \in \{\alpha, \beta\}$and his position in the order of opportunities to withdraw $l \in [0,1]$:

[5]Taxation with the purpose of funding the deposit insurance plan of the government has been introduced into the Diamond and Dybvig framework by Freeman (1998), Boyd et al. (2002) and Martin (2006). In this paper, following Keister (2016), the government's tax revenues serve a more general purpose since they are used to provide a public good, in addition to bailing out the financial sector.

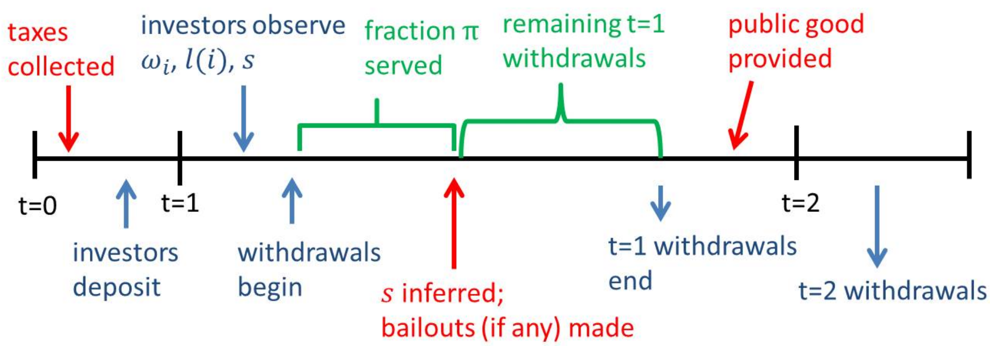


Figure 1.1: Timeline

$$y_i : \{0,1\} \times \{\alpha, \beta\} \times [0,1] \to \{0,1\}$$

where $y_i = 0$ corresponds to withdrawing in period 1 and $y_i = 1$ corresponds to withdrawing in period 2. On the other hand, the banks and the government do not observe the realization of the sunspot state. Nevertheless, we will see that in equilibrium banks are able to infer the realization of s after the measure of withdrawals exceeds a certain threshold.

1.2.3 Timeline

At the start of period 0, investors receive their initial endowment and the government collects taxes by imposing a common tax rate τ on investors' endowments. Investors then deposit their after-tax endowment in the banking sector and period 0 ends.

At the beginning of period 1, each investor $i \in [0,1]$ observes whether he is patient or impatient $\omega_i \in \{0,1\}$, his position in the order of opportunities to withdraw $l(i) \in [0,1]$ and the realization of the sunspot state $s \in \{\alpha, \beta\}$. Withdrawals then begin. The environment here is similar to Ennis and Keister (2010). Specifically, each impatient investor always strictly prefers to withdraw in period 1, regardless of the realization of s or his position in the order of opportunities to withdraw. Moreover, given that the fraction of impatient investors in period 1 is always equal to π, banks are unable to infer the realization of s while the first π fraction of withdrawals is being made. Hence, in equilibrium, payments to the first π fraction of investors to withdraw cannot be contingent on s. After π withdrawals, banks will be able to infer the state. In particular, if the state is α, there are no more withdrawals in period 1 since all of the impatient investors were able to contact the bank and withdraw. In contrast, if the state is β and a run is underway, withdrawals will continue because some impatient investors have not yet been able to withdraw. The government, in this case, has the option of using fraction of the tax revenues in order to make transfers to those (and only those) banks experiencing a run. After any bailouts have been made, the remaining tax revenue is used to provide the public good and period 1 ends.[6]

In period 2, those patient investors that did withdraw in period 1 receive a pro-rata share of the bank's resources in the final period and the game ends.

1.3 Panic equilibrium

Given the self-fulfilling nature of a run in this model, a no-panic equilibrium, where investors withdraw in period 1 only when they are impatient, always exists. At the same time, another equilibrium where fraction of the patient investors panic and withdraw in period 1 for certain realizations of the sunspot state may also exist. Equilibrium of the later type will be called a *panic equilibrium* and will be the focus of this section.

[6] Observe that the central location in period 1 will be contacted only by those demanding to withdraw in period 1, which, in addition to the fact that there is no aggregate uncertainty about the fraction of truly impatient investors, allows the bank to completely infer the state after the measure of withdrawals reaches π. Notice that the approach here is different from Green and Lin (2003), where all investors must report to the bank in period 1.

1.3.1 A profile of withdrawal strategies

I focus on equilibria where investors with the same endowment follow the same strategy and I introduce two strategies – the *no-panic* and the *panic* strategy. The *no-panic strategy* is the standard truth-telling behavior in a Diamond and Dybvig model, that is, investors with endowment e are said to follow the *no-panic strategy* if they choose to withdraw in period 1 only when impatient.

$$y_i^{NP}(\omega_i, \alpha, l) = \omega_i \text{ for } i \in [0,1] \text{ s.t. } E(i) = e \tag{1.1}$$

On the other hand, investors with endowment e are said to follow the *panic strategy* if:

$$\left\{ \begin{array}{c} y_i^{P}(\omega_i, \alpha, l) = \omega_i \\ y_i^{P}(\omega_i, \beta, l) = \left\{ \begin{array}{c} 0 \\ \omega_i \end{array} \right\} \text{ if } \left\{ \begin{array}{c} l \leq \pi \\ l > \pi \end{array} \right\} \end{array} \right\} \quad \text{for} \quad i \in [0,1] \text{ s.t. } E(i) = e \tag{1.2}$$

Impatient investors always withdraw in period 1 because they do not derive any utility from consuming in the last period. Patient investors, on the other hand, face a strategic choice. According to (1.2) when the state is α, patient investors with endowment e choose to wait until period 2 to withdraw. In contrast, when the state is β, patient investors with endowment e choose to withdraw in period 1 when they are among the first π fraction of investors with an opportunity to withdraw.[7] Observe that the strategy profile in (1.2) specifies that the run stops after π withdrawals have taken place and the sunspot state is inferred by the banks. This type of strategy

[7] The panic strategy in (1.2) assumes that each investor knows his exact order in the opportunities to withdraw $l \in [0,1]$ during period 1 and this information is used when deciding whether to withdraw in period 1 or wait until period 2. However, all results that will be presented obtain under the weaker assumption according to which each investor only knows whether he is able to withdraw before banks infer the realization of the sunspot.

profile was introduced by Ennis and Keister (2010), who showed that in settings where banks are able to react by changing payments when withdrawal demand is high an equilibrium bank run will be necessarily partial and restricted to those investors that can withdraw before banks infer the state.

The government's choice of τ at the beginning of period 0 leads to a proper sub-game associated with this value of the tax rate. Lack of commitment implies that the actions of banks and the government within this sub-game are taken after investors have chosen withdrawal strategies and, therefore, must be a best response to those strategies. It will be convenient to characterize the profile of withdrawal strategies by defining $P(\tau)$ as the *set of investor types* that follow the panic strategy in the *sub-game for given* τ. For example, a panic set of the form $P(\tau) = [x, e_H]$ means that investors with endowment in the set $[x, e_H]$ follow the panic strategy, whereas the remaining investor types $[e_L, x)$ follow the no-panic strategy.

1.3.2 Type-specific banks

To begin, I will study equilibrium of the model assuming that investors with the same endowment operate their own liquidity insurance arrangement (i.e. a separate bank for type e investors).

Definition 1.1. The part of the financial system providing intermediation only to type e investors is called *type e banks*

Assuming that each type of investor operates its own bank simplifies the analysis. In particular, the intervention of the government can be characterized in terms of a bailout to type e banks, or equivalently, in terms of a bailout to type e investors since there is a *one-to-one correspondence* between the type of investors and their banks. In reality, financial intermediaries perform maturity transformation for investors with different wealth levels. Notice, however, that "a bank" in this framework should not be interpreted in the usual sense as a separate legal entity, but rather as an equilibrium clustering of investors of the same type within the financial system. In section 1.5, the model is augmented to include a bank formation stage in period 0 and I show that, under plausible assumptions about the behavior of financial intermediaries, the structure of the financial sector emerging in equilibrium is equivalent to assuming a separate bank for each type of investors. For this reason, I first present the analysis in the simpler case where this pattern is simply assumed.

The payment schedule for type e investors in the subgame for τ is summarized in the following vector:

$$\left(c_1\left(e,\tau\right), c_{2\alpha}\left(e,\tau\right), c_{1\beta}\left(e,\tau\right), c_{2\beta}\left(e,\tau\right)\right) \tag{1.3}$$

The payment given to investors withdrawing before banks infer the state $c_1\left(e,\tau\right)$ cannot be contingent on s. After fraction π of withdrawal has taken place, banks infer the state s and reschedule payments in order to reflect this new information. If the state is α, all of the remaining investors are patient and each receives $c_{2\alpha}\left(e,\tau\right)$ in period 2. On the other hand, if the state is β, each of the remaining impatient investors receives $c_{1\beta}\left(e,\tau\right)$ in period 1, whereas each of the remaining patient investors receives $c_{2\beta}\left(e,\tau\right)$ in period 2.

Importantly the payment schedule in (1.3) is not chosen ex-ante, but rather will be determined as the outcome of a process in which banks updates their payments in order to reflect the arrival of new information in an ex-post optimal way. Given the payment schedule in (1.3), type e investors will be best responding with the panic strategy in (1.2) whenever the following set of conditions are satisfied:

$$\frac{c_1(e,\tau)}{c_{2\alpha}(e,\tau)} < 1 \quad \textsf{and} \quad \frac{c_1(e,\tau)}{c_{2\beta}(e,\tau)} > 1 \tag{1.4}$$

To see that (1.4) is sufficient for type e investors to best respond with the panic strategy in (1.2), consider a patient type e investor with a chance to withdraw before s is inferred by the banks. If $s = \alpha$, the first inequality in (1.4) ensures that he prefers to wait. On the other hand, if $s = \beta$, the second inequality in (1.4) ensures that he prefers to withdraw in period 1.

I search for equilibria of the sub-game for given τ by: (i) fixing a panic set of investor types $P(\tau) \subseteq [e_L, e_H]$, (ii) deriving the best response of the banks and the government to this complete profile of withdrawal strategies, (iii) checking whether condition (1.4) is satisfied for each investor type in the panic set $P(\tau)$, and, finally, (iv) finding the tax rate τ^* in period 0 that yields the highest aggregate expected utility.[8]

1.3.3 Withdrawals after banks infer s

If type e investors follow the no-panic strategy, then after π withdrawals have been made in period 1, all of the remaining investors will be patient, regardless of the state s. If we define $\pi_s(e)$ to be the fraction of the remaining type e investors who are impatient, we have:

$$\pi_\alpha(e) = 0 \quad and \quad \pi_\beta(e) = 0 \tag{1.5}$$

On the other hand, if type e investors follow the panic strategy, we have:

$$\pi_\alpha(e) = 0 \quad and \quad \pi_\beta(e) = \pi \tag{1.6}$$

If withdrawals stop after reaching a measure of π, the bank infers that the state is α and therefore the first π withdrawals were made only by impatient type e investors, hence $\pi_\alpha(e) = 0$. If withdrawals continue after π, the bank infers that the state is β and the fraction of the remaining type e investors who are impatient equals π.[9] Let

[8]Given that the run on each bank is sustained by self-fulfilling expectations, we can always find equilibria where all investors in a given bank follow the no-panic strategy. As a result, investor types outside the panic set will be best responding with the no-panic strategy.

[9]Patient and impatient investors are equally likely to occupy any place in the order of opportunities to withdraw. When the state is β and after π withdrawals, the fraction of impatient investors who have been served will be equal to π^2, which implies that the remaining fraction of impatient

$\psi_s(e,\tau)$ denote the quantity of per capita resources after the bank have serviced a fraction π of the investors. If $s=\alpha$:

$$\psi_\alpha(e,\tau) = (1-\tau)e - \pi c_1(e,\tau) \tag{1.7}$$

On the other hand, if $s=\beta$ the bank's resources can be potentially augmented by a bailout transfer from the government. Letting $b(e,\tau)$ denote the per capita bailout to a type e bank, we have:

$$\psi_\beta(e,\tau) = (1-\tau)e - \pi c_1(e,\tau) + b(e,\tau) \tag{1.8}$$

The remaining resources of the bank in state s will be distributed efficiently among the remaining investors:

$$V\left(\psi_s(e,\tau);\, \pi_s(e)\right) = \max_{c_{1s}(e,\tau),c_{2s}(e,\tau)} (1-\pi)\left[\pi_s(e)u\left(c_{1s}(e,\tau)\right) + \left(1-\pi_s(e)\right)u\left(c_{2s}(e,\tau)\right)\right] \tag{1.9}$$

subject to the budget constraint in state s:

$$(1-\pi)\left[\pi_s(e)c_{1s}(e,\tau) + \left(1-\pi_s(e)\right)\frac{c_{2s}(e,\tau)}{R}\right] = \psi_s(e,\tau) \tag{1.10}$$

where $\pi_s(e)$ is determined by (1.5) if type e follows the no-panic strategy and by (1.6) if type e follows the panic strategy. The bank's payments after inferring the state will satisfy the following first order condition, where $\mu_s(e,\tau)$ is the shadow value on the resources constraint in state s:

$$\mu_s(e,\tau) = u'(c_{1s}(e,\tau)) = Ru'(c_{2s}(e,\tau)) \tag{1.11}$$

The above condition implies that waiting to withdraw in period 2 becomes a dominant strategy for patient investors once banks infer the state and reschedule payments. In other words, as in Ennis and Keister (2010), an equilibrium bank run in this setting is necessarily partial and can only involve investors who are able to withdraw before

investors will be equal to $\dfrac{\pi-\pi^2}{1-\pi} = \pi$.

banks infer the state.

1.3.4 Bailouts

If $s = \alpha$ the government does not intervene in the banking sector and the level of the public good is equal to the tax proceeds $g_\alpha(\tau) = \tau$. On the other hand, if $s = \beta$ the government allocates the tax proceeds between the public good and bailouts to banks experiencing a run in order to maximize:

$$\underset{b(e,\tau)\geq 0}{max} \int_{e\in P(\tau)} V\left((1-\tau)e - \pi c_1(e,\tau) + b(e,\tau);\, \pi\right) dG(e) + v\left(\tau - B(\tau)\right) \tag{1.12}$$

The bailout transfer must be non-negative (that is, taxing after period 0 is not allowed) and bailouts will be potentially made only to those banks whose investors follow the panic strategy (i.e. $e \in P(\tau)$), since these are the banks that experience a run when $s = \beta$. The public good in state β is equal to the remaining tax revenues $g_\beta(\tau) = \tau - B(\tau)$, where $B(\tau)$ is the aggregate bailout:

$$B(\tau) = \int_{e\in P(\tau)} b(e,\tau) dG(e) \tag{1.13}$$

Differentiating (1.12) with respect to $b(e,\tau)$ and taking into account that bailouts must be non-negative, we obtain:

$$\mu_\beta(e,\tau) \leq v'(\tau - B(\tau)) \tag{1.14}$$

where (1.14) holds with equality whenever $b(e,\tau) > 0$. According to (1.14), the marginal utility from the private good will be equalized to the marginal utility from the public good in all banks receiving a bailout. In addition, (1.11) and (1.14) imply that all banks receiving a bailout will provide the same payment schedule to their remaining investors in state β, which will be independent of e and henceforth denoted $(c_{1\beta}(\tau), c_{2\beta}(\tau))$. From (1.6) and (1.10) we obtain that $\psi_\beta^B(\tau)$ - the per-capita level of the resources necessary to deliver this payment schedule - will satisfy:

$$(1-\pi)\left[\pi c_{1\beta}(\tau) + (1-\pi)\,\frac{c_{2\beta}(\tau)}{R}\right] = \psi_{\beta}^{B}(\tau) \tag{1.15}$$

Observe that (1.14) and (1.15) imply that $\psi_{\beta}^{B}(\tau)$ is determined by aggregate conditions and is, therefore, treated as exogenous by individual banks. Finally, the bailout transfer to type e banks (when their investors follow the panic strategy) is characterized by:

$$b(e,\tau) = max\left\{\psi_{\beta}^{B}(\tau) - [(1-\tau)e - \pi c_1(e,\tau)]\,,\, 0\right\} \tag{1.16}$$

The next result shows that the bailout to a given bank is decreasing in the endowment of its investors:

Proposition 1.2. ***(The ex-post optimal bailout policy)*** *Suppose that type e_1 and type e_2 investors with $e_1 < e_2$ follow the panic strategy in equilibrium. Then, if either type e_1 or type e_2 bank receive a bailout, we have $b(e_1,\tau) > b(e_2,\tau)$*

The government's intervention ensures that the payment schedule is the same across all banks that have been bailed out. In equilibrium, banks whose investors have higher initial endowment receive lower bailouts (if any at all), since these banks would also have a higher level of per capita resources before the government's intervention and therefore require less in government transfers in order to deliver the common payment schedule in banks receiving a bailout.

1.3.5 Withdrawals before banks infer s

During the first π withdrawals, banks are uncertain about the realization of the state s and therefore payments cannot be made contingent s. In addition, banks whose investors follow the panic strategy will experience a run when $s = \beta$ and therefore could *qualify* for a bailout from the government in that state. From (1.16), a bank experiencing a run in state β will be bailed out by the government whenever the quantity of this bank's per-capita resources is below $\psi^B_\beta(\tau)$. Thus, depending on the choice of $c_1(e,\tau)$, a type e bank qualifies for a bailout if:

$$(1-\tau)e - \pi c_1(e,\tau) < \psi^B_\beta(\tau) \tag{1.17}$$

and does not qualify if:

$$(1-\tau)e - \pi c_1(e,\tau) \geq \psi^B_\beta(\tau) \tag{1.18}$$

Suppose that type e investors follow the panic strategy. A type e bank that *chooses* to qualify for a bailout will solve the following program:

$$W^B(e,\tau) \equiv \max_{c_1(e,\tau)} \left\{\pi u\left(c_1(e,\tau)\right) + (1-q)V\left((1-\tau)e - \pi c_1(e,\tau);\, \pi_\alpha\right) + qV\left(\psi^B_\beta(\tau);\, \pi_\beta\right)\right\} \tag{1.19}$$

subject to (1.17), which ensures that its choice of $c_1(e,\tau)$ will prompt the government to bail out the bank in state β. On the other hand, a type e bank that chooses not to qualify for a bailout will solve the following program:

$$W^{NB}(e,\tau) \equiv \max_{c_1(e,\tau)} \left\{ \begin{array}{c} \pi u\left(c_1(e,\tau)\right) + (1-q)V\left((1-\tau)e - \pi c_1(e,\tau);\, \pi_\alpha\right) \\ +qV\left((1-\tau)e - \pi c_1(e,\tau);\, \pi_\beta\right) \end{array} \right\} \tag{1.20}$$

subject to (1.18), which ensures that its choice of $c_1(e,\tau)$ will not prompt the government to bail out the bank in state β. Henceforth, a bank that solves the program in (1.20) is said to self-insure. A fraction π of type e investors withdraw before the bank has inferred the state and each one of them receives $c_1(e,\tau)$.[10] The second and third terms are characterized by (1.9) and represent the expected utility for the remaining investors in the bank after the state is inferred and the remaining resources

[10] Investors are risk averse and therefore every type e bank chooses to give the same payment $c_1(e,\tau)$ to all investors withdrawing before s is inferred.

are distributed efficiently. The difference between (1.19) and (1.20) is that banks receiving a bailout in state β will have the same quantity of per-capita resources in this state $\psi_\beta^B(\tau)$, which will be determined by economy-wide conditions. Next, define the function:

$$D(e,\tau) \equiv W^B(e,\tau) - W^{NB}(e,\tau) \tag{1.21}$$

If $D(e,\tau) > 0$, type e banks best respond by qualifying for a bailout, whereas, if $D(e,\tau) < 0$, type e banks best respond by self-insuring. Denoting with $c_1^B(e,\tau)$ the solution to the program in (1.19) and with $c_1^{NB}(e,\tau)$ the solution to the program in (1.20). The choice of payments during the first π withdrawals in banks whose investors follow the panic strategy is characterized below:

Proposition 1.3. *In a sub-game for given τ:*

(i) The choice of $c_1(e,\tau)$ when type e banks best respond by qualifying for a bailout is characterized by:

$$u'(c_1^B(e,\tau)) = (1-q)\mu_\alpha^B(e,\tau) \tag{1.22}$$

(ii) The choice of $c_1(e,\tau)$ when type e banks best respond by self-insuring is characterized by:

$$u'(c_1^{NB}(e,\tau)) = (1-q)\mu_\alpha^{NB}(e,\tau) + q\mu_\beta^{NB}(e,\tau) \tag{1.23}$$

A bank that qualifies for a bailout will have the same quantity of per-capita resources in state β, regardless of its choice of payments to the first π investors (as long as its choice satisfies (1.17)). Hence, a bank that qualifies for a bailout ignores the shadow value of its resources in state β and sets early payments according to (1.22). On the other hand, a bank that self-insures takes into account that higher $c_1(e,\tau)$ leads to a lower level of resources both when the state is α and when the state is β and sets the payments to the first π investors according to (1.23), which would ensure that the marginal utility of investors withdrawing before the state is revealed equals the expected shadow value of the bank's resources after the state is known. Henceforth, I restrict the parameters of the model to satisfy:

$$q < \frac{R-1}{R} \tag{1.24}$$

$$\frac{R^{\frac{1-\gamma}{\gamma}}}{\pi + (1-\pi)R^{\frac{1-\gamma}{\gamma}}} < \left(\frac{1-Rq}{R-Rq}\right)^{\frac{1}{\gamma}} \tag{1.25}$$

Lemma 1 in the appendix shows that (1.24) is necessary for the the existence of equilibria where investors in banks qualifying for a bailout best respond with the panic strategy in (1.2). At the same time, investors in banks that self-insure will best respond with the panic strategy whenever the model parameters satisfy (1.25).[11] Next, we can show that if type e banks best respond by qualifying for a bailout, they would also increase the payments to investors withdrawing before s is known.

Proposition 1.4. *For all τ and e, we have $c_1^B(e,\tau) > c_1^{NB}(e,\tau)$*

Furthermore, whenever a bank finds it optimal to qualify for a bailout, the increase in the payments to investors withdrawing before s is known will be independent of the actual size of the bailout. This fact can be seen from combining (1.22) and (1.23) in order to obtain:

$$\frac{u'(c_1^{NB}(e,\tau))}{u'(c_1^B(e,\tau))} = \frac{\mu_\alpha^{NB}(e,\tau)}{\mu_\alpha^B(e,\tau)} + \frac{q}{(1-q)}\frac{\mu_\beta^{NB}(e,\tau)}{\mu_\alpha^B(e,\tau)}$$

Observe from (1.7), (1.10) and (1.11) that none of the variables on the right hand side is a function of the bailout to type e banks. Hence, the increase in $c_1^B(e,\tau)$ relative to $c_1^{NB}(e,\tau)$ will be independent of $b(e,\tau)$. The implication is that a bank must anticipate a minimum level of a bailout in order to be willing to choose $c_1^B(e,\tau)$

[11] Condition (1.24) assumes that the probability of the panic state is not too high and is identical to the one in Keister (2015). If (1.24) does not hold, then the panic strategy in (1.2)(1.2) cannot be consistent with equilibrium because investors in banks that are being bailed out in state β will have an incentive to run regardless of the realization of s.

and hence incur the subsequent reduction in the payments to the remaining investors when the state turns out to be α and there is no run. In other words, if the bailout payment the bank would receive in state β is relatively small, the bank would prefer to self-insure by choosing $c_1^{NB}(e,\tau)$. Next , define the function $e^{NB}(\tau)$:

$$e^{NB}(\tau) \equiv \left\{ \begin{array}{c} e_L \\ z \in (e_L, e_H) \;\; s.t \; D(z,\tau) = 0 \\ e_H \end{array} \right\} \quad if \quad \left\{ \begin{array}{c} D(e_L,\tau) \leq 0 \\ D(e_L,\tau) > 0 \; and \; D(e_H,\tau) < 0 \\ D(e_H,\tau) \geq 0 \end{array} \right\} \tag{1.26}$$

We can show the following:

Proposition 1.5. *(**Banks best response for given** τ) In the sub-game for a particular value of τ, let $e^{NB}(\tau)$ be given by* (1.26) *and consider banks whose investors follow the panic strategy in* (1.2)*.*

(i) if $e^{NB}(\tau) = e_L$, then banks would choose to self-insure

(iii) if $e^{NB}(\tau) \in (e_L, e_H)$, then banks with $e < e^{NB}(\tau)$ choose to qualify for a bailout, whereas banks with $e > e^{NB}(\tau)$ choose to self-insure.

(iii) if $e^{NB}(\tau) = e_H$, then all banks would choose to qualify for a bailout

In order to see why this result holds, take the derivative of $D(e,\tau)$ with respect to e, apply the envelope theorem, and use $c_1^B(e,\tau) > c_1^{NB}(e,\tau)$ to obtain:

$$\frac{\partial D(e,\tau)}{\partial e} = (1-\tau)\left(u'(c_1^B(e,\tau)) - u'(c_1^{NB}(e,\tau))\right) < 0 \tag{1.27}$$

hence $D(e,\tau)$ is a strictly decreasing function of e and we obtain the result in Proposition 1.5. Ex-ante, each bank has the option to self-insure by solving the program in (1.20). The desirability of self-insurance will be higher for banks whose investors have higher initial endowment, since they anticipate to receive lower bailouts from the government in case of a run. In addition, we can show that:

$$\frac{\partial e^{NB}(\tau)}{\partial \tau} > 0$$

For larger values of τ, banks would anticipate larger bailouts and, as a result, larger fraction of them are willing to choose payments which qualify them for a bailout in case they experience a run. That is, the option to self-insure by choosing $c_1^{NB}(e, \tau)$ becomes less desirable for a type e bank when the government is expected to provide more generous bailouts in a panic.

Discussion

The analysis in this section, specifically the idea that any given bank must decide whether to take actions which will make it *eligible* for a bailout in the event of a run from its investors, is novel to the literature on financial panics. Moreover, it might initially appear that it can never be optimal for any bank to disqualify itself for a bailout, since a bailout is an inflow of resources, which allows the bank to provide higher payments to its remaining investors. Indeed, for fixed $c_1(e, \tau)$, a type e bank will be able to deliver higher ex-ante utility to its investors for any $b(e, \tau) > 0$. The crucial observation, however, is that the choice of $c_1(e, \tau)$ is affected by the decision of the bank to qualify (or disqualify) itself for a bailout. That is, once a given bank chooses to follow a strategy which secures a bailout in case of a run, it also becomes optimal to aggressively increase payments when there is uncertainty as to whether the run will actually take place. The negative side of obtaining a bailout is precisely this discrete jump in the payments to investors withdrawing before the state is known, which depletes the resources of the bank and thus imposes a cost to the remaining investors when the state is α and the run does not take place.

In addition, while it might be optimal for individual banks to receive a bailout, the cost of the government intervention will be too high from a social perspective. The reason is that banks fail to *internalize* their collective effect on the level of the public good in state β and end up choosing payments which are too high (from the perspective of a social planner) which would increase the cost of a subsequent government intervention during a panic.[12]

1.3.6 Equilibria for fixed τ

For a particular value of τ, consider a complete profile of withdrawal strategies characterized by a panic set $P(\tau) \subseteq [e_L, e_L]$. Each investor type e in the panic set $P(\tau)$ follows the panic strategy in (1.2) whereas each investor type not in the panic set follows the no-panic strategy in (1.1). A panic equilibrium is associated with a non-empty panic set of investor types $P(\tau) \neq \emptyset$.

The payment schedule for each investor type that follows the panic strategy is characterized by (1.6), (1.10), (1.11), and (1.19) if their bank receives a bailout and by (1.6), (1.10), (1.11) and (1.20) if it does not. The government bailout policy is characterized by (1.13) - (1.16). The payment schedule for type e investors in case they follow the no-panic strategy is characterized by (1.5), (1.10), (1.11), and (1.23). Finally, condition (1.4) must hold for each investor type that follows the panic strategy, since this would ensure that they are best responding with the panic strategy in (1.2).

In this section, the dependence of a given equilibrium variable on the panic set of investors is denoted explicitly. That is, I write $z\left(e, \tau; P(\tau)\right)$ whenever the equilibrium value of z is a function of τ, e and $P(\tau)$. From, (1.10), (1.11), (1.22) and (1.23) we obtain that payments in banks that do not receive a bailout in state β - either because their investors follow the no-panic strategy or because these banks best respond by self-insuring instead of qualifying for a bailout – do not depend on $P(\tau)$. In fact, the only component of the banks' payments schedule that depends directly on the panic set $P(\tau)$ is the allocation in state β in banks that qualify for a bailout i.e.

[12] Keister (2016) analyzes this distortion and policies that aim to correct it in a model in which all investors are identical.

$\left(c_{1\beta}^{B}\left(\tau;P(\tau)\right),c_{2\beta}^{B}\left(\tau;P(\tau)\right)\right)$.

Observe that having a fraction of the investors follow the panic strategy constitutes a necessary, but not a sufficient condition for bailouts to be made in equilibrium, since, as we saw in the previous section, each banks will choose whether to qualify for a bailout in the event of a run or, instead, to self-insure. That is, whether or not the equilibrium run on the banks leads to a bailout intervention is an endogenous outcome of the model. So fix τ and consider an investor type x and a panic set $P=[x,e_H]$ such that the following holds:

$$\left\{\begin{array}{c} u'\left(c_{1\beta}^{B}\left(\tau;[x,e_H]\right)\right)=v'\left(\tau-\int_{e\in[x,e_H]}b\left(e,\tau;[x,e_H]\right)dG\right) \\ D\left(x,\tau;[x,e_H]\right)>0 \\ c_{1}^{B}\left(x,\tau\right)\geq c_{2\beta}^{B}\left(\tau;[x,e_H]\right) \end{array}\right\} \tag{1.28}$$

The equation on the first line in (1.28) implies that the bailout policy of the government is ex-post optimal when investors with endowment in the set $[x,e_H]$ follow the panic strategy. The inequality on the second line implies that type x banks best respond by qualifying for a bailout, whereas inequality on the third line implies that investors whose endowment is x and who are able to withdraw before banks infer the state would best respond by doing so. Let $X\left(\tau\right)$ denote the set of all investor types for which (1.28) is satisfied in the sub-game for τ:

$$X(\tau)\equiv\{x\in(e_L,e_H)\ \text{s.t. (28) holds}\} \tag{1.29}$$

The next proposition shows that each investor type in $X(\tau)$ can be used to construct a panic equilibrium that involves bailouts in state β:

Proposition 1.6. ***(Panic equilibrium with bailouts)*** *If $x\in X(\tau)$, then there is a panic equilibrium in the sub-game for τ where investors with endowment in $[e_L,x)$ best respond with the no-panic strategy and investors with endowment in $[x,e_H]$ best respond with the panic strategy. Moreover, this equilibrium involves bailouts when*

$s = \beta$.

On the other hand, if $X(\tau)$ is empty, then for each $x \in (e_L, e_H)$: either type x banks best respond by choosing to self-insure instead of qualifying for a bailout or type x banks best respond by qualifying for a bailout, but their investors face no incentive to panic in state β. Nevertheless, a panic equilibrium may still exist provided that we can find a panic set of investor types $P(\tau) = [y, e_H]$ such that the following is satisfied:

$$\left\{ \begin{array}{c} u'\left(c_{1\beta}^{NB}(y,\tau)\right) \leq v'(\tau) \\ D\left(y, \tau; [y, e_H]\right) \leq 0 \end{array} \right\} \tag{1.30}$$

The first inequality in (1.30) implies that the government would not bail out type y banks when they choose to self-insure, whereas, the second inequality implies that type y banks best respond by self-insuring. Next, define $Y(\tau)$ to be the set of all investor types for which (1.30) is satisfied:

$$Y(\tau) \equiv \{y \in (e_L, e_H) \text{ s.t. (30) holds}\} \tag{1.31}$$

The next proposition shows that each element in the set $Y(\tau)$ can be associated with a panic equilibrium such that no-bailouts are made in state β.

Proposition 1.7. *(**Panic equilibrium without bailouts**) For each $y \in Y(\tau)$, there exists an equilibrium where investors with endowment $e \in [e_L, y)$ follow the no-panic strategy and investors with endowment $e \in [y, e_H]$ follow the panic strategy. There will be no bailouts in equilibrium.*

One question that arises in this setting is the following: which investor types could form a part of an equilibrium run on the banks for a given τ? In order to make this question more interesting, I will assume that the panic set $[e_L, e_H]$ is not consistent with equilibrium in the subgame for τ (otherwise, the answer is readily available since the equilibrium where all investors follow the panic strategy will exist). In order to find which investors cannot be part of a run on the banks I need to introduce some additional notation. Define the set of *threshold investor types* for a given τ:

$$T(\tau) \equiv \left\{x \in X(\tau) \text{ } and \text{ } c_1^B(x,\tau) = c_{2\beta}^B(\tau;[x,e_H])\right\} \tag{1.32}$$

Investor type belonging to $X(\tau)$ and that is also indifferent, in state β, between withdrawing in period 1 before s is inferred and waiting until period 2 to withdraw will be called a threshold investor type. If $T(\tau)$ is non-empty, let $e^T(\tau)$ denote the lowest among the threshold endowment types for a given τ:[13]

$$e^T(\tau) \equiv \{x \in T(\tau) \text{ } s.t \text{ } x \leq x' \text{ } for \text{ } each \text{ } x' \in T(\tau)\} \tag{1.33}$$

We are now ready to characterize the type of investors that could participate in an equilibrium run on the banks for a given value of τ.

Proposition 1.8. *For a fixed τ, suppose that (i) $[e_L, e_H]$ is not an equilibrium panic set and (ii) $T(\tau)$ is non-empty. Then an investor type whose endowment belongs to the interval $[e_L, e^T(\tau))$ cannot be part of an equilibrium run on the banks.*

The government's bailout policy deters investors whose endowment is sufficiently low from running on the banks as part of equilibrium in the sub-game for τ. Moreover, since investors with endowment in $[e_L, e^T(\tau))$ do not panic, there will be no need to bail them out as part of equilibrium. Thus, the bailout policy with respect to these types of investors is similar to deposit insurance in the standard Diamond and Dybvig model where an off-equilibrium promise is enough to eliminate the run at no cost in equilibrium. Unlike in the standard case, however, both the fraction of the investors covered by deposit insurance and their characteristics (i.e. those whose endowment

[13]The condition for the uniqueness of the threshold investor type depends in a complicated way on the parameters of the model. Nevertheless, the uniqueness of the threshold investor type is not crucial for the the main insights of the model.

is sufficiently low) will be determined endogenously.

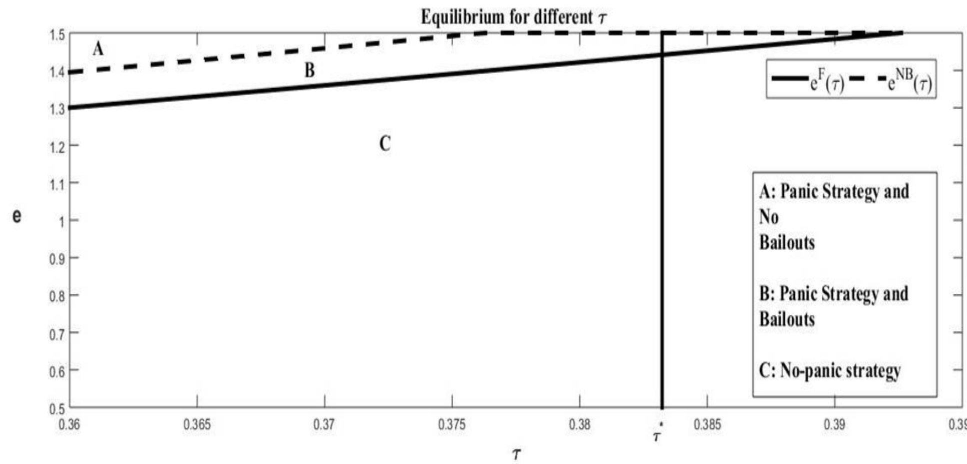
At the same time, a promise of more generous rescue package for investors whose endowment exceeds $e^T(\tau)$ in order to deter them from panicking fails to be credible since investors correctly perceive that such a rescue package will not be ex-post optimal for the government. As a result, an equilibrium run on the banks could engulf investors whose endowment exceeds $e^T(\tau)$. The last result also shows that a panic equilibrium where investors follow the panic strategy only if their endowment is above a certain threshold was not selected arbitrarily, but instead would arise naturally once we account for the way the government's bailout policy would influences investors' propensity to run on the banks.

Numerical example

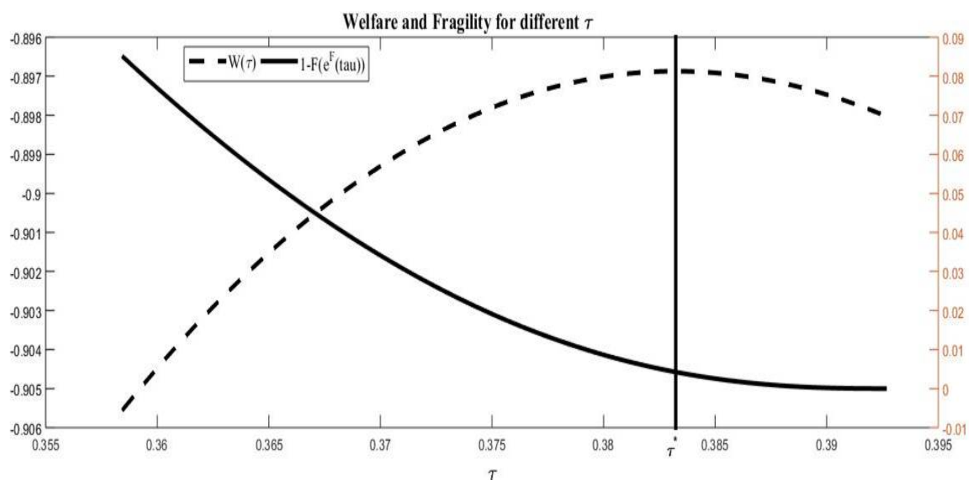
Panic equilibria characterized by a threshold investor type are shown for different values of τ in panel (a) of figure 1.2. The tax rate τ is plotted on the horizontal axes and investors' endowments are plotted on the vertical axes. Endowments are generated by $e = 0.5 + x$, where x follows a symmetric beta distribution. The remaining parameters of the model are set as follows: $R = 3$, $\pi = 0.5$, $\gamma = 5$, $\delta = 1$ and $q = 0.05$.[14] The threshold investor type $e^T(\tau)$ exists and is unique for the range of τ depicted on the figure. Moreover, panel (b) shows that the tax rate that will, in fact, be chosen by the government also belongs to this range (more on the choice of τ in the next section). Region A in the figure is between $e^{NB}(\tau)$ and e_H and it depicts combinations of (e, τ) such that type e investors follow the panic strategy *and* type e banks are not bailed out in state β. Region B is between $e^T(\tau)$ and $e^{NB}(\tau)$ and it depicts combinations of (e, τ) such that type e investors follow the panic strategy and type e banks receive a bailout in state β. Region C is between e_L and $e^T(\tau)$ and it shows combinations of (e, τ) for which type e investors necessarily follow the no-panic

[14]Numerical investigations suggest that other plausible specifications for the distribution for investors' endowments, like the normal or the exponential distribution yield similar qualitative results.

strategy in equilibrium. Observe that higher values of τ also lead to larger bailouts in a crisis and, as a result, the fraction of investors for which the panic strategy cannot be a best response and the fraction of banks choosing to qualify for a bailout instead of self-insuring are both increasing in τ (i.e $e^T(\tau)$ and $e^{NB}(\tau)$ are increasing in τ).



(a) Panic equilibria for fixed τ



(b) The choice of τ

Figure 1.2: Equilibrium for a given tax rate

1.3.7 The choice of τ

For given τ, panel (b) in figure 1.2 plots the aggregate welfare $W(\tau)$ associated with the panic equilibrium on panel (a).

$$W(\tau) = \int_{e_L}^{e^T(\tau)} W^{NP}(e,\tau)dG + \int_{e^T(\tau)}^{e^{NB}(\tau)} W^{B}(e,\tau)dG + \int_{e^{NB}(\tau)}^{e_H} W^{NB}(e,\tau)dG + Ev(g_s(\tau)) \tag{1.34}$$

where last term is the expected utility from the public good:

$$Ev(g_s(\tau)) = (1-q)v(\tau) + qv(\tau - B(\tau)) \tag{1.35}$$

Panel (b) also plots the fraction of investors that are best responding with the panic strategy for given a value of the tax rate $(1 - G\left(e^T(\tau)\right))$. The objective of the government in period 0 is to choose the tax rate associated with the maximum welfare in the resulting sub-game. Note that the government is facing a trade-off between protecting the economy from a crisis and efficiently allocation resources during non-crisis times. If the probability of a crisis were zero (i.e. $q = 0$), the government collects taxes only with the goal of providing the public good. On the other hand, if $q > 0$ then an additional dollar in tax revenues become more valuable since it can be used to bail out the banks in addition to providing a public good. At the same time, increasing τ would lead to inefficiently high tax burden when the state is α and the run does not take place. The tax rate chosen by the government τ^* is depicted by the vertical line in both panels and it allows for a fraction of the investors to be best responding with the panic strategy. We have established the following proposition:

Proposition 1.9. *For some parameter values there exists an equilibrium in which a positive fraction of investors follow the panic strategy and bailouts are made to all banks experiencing a run*

Hence, we can have an equilibrium in the overall game where some investors panic and run on the banks. Moreover, if such an equilibrium exists, the panic and the subsequent bailouts will be restricted to investors that are relatively wealthy. At the same time, by depressing the level of the public good in state β, the cost of the bailout would be affecting all agents in the economy.

1.4 Inequality and financial fragility

In this section, the distribution of investors' endowments is restricted to have a two point support $\{e_L, e_H\}$, where e_H is the endowment of the high-income investors and e_L is the endowment of the low-income investors. Investors' endowment are generated in the following way:

$$e_L(\Delta) = 1 - f\Delta \quad and \quad e_H(\Delta) = 1 + (1-f)\Delta \tag{1.36}$$

where $f \in (0,1)$ denotes the fraction of high-income investors and where the level of inequality is captured by the wealth gap $\Delta \in [0,\, f^{-1})$ between the high and low income investors:

$$\Delta = e_H - e_L \tag{1.37}$$

Hence, the average endowment is always 1, there is no wealth inequality when $\Delta = 0$ and higher values of Δ are associated with more unequal wealth distribution. An economy will be characterized by the following vector of parameter values R, π, γ, δ, q, f and Δ.[15]

1.4.1 Inequality and fragility for fixed τ

I begin by analyzing the effect of inequality on the nature of financial fragility for a fixed value of τ.

[15] Also, note that higher values of Δ are associated with a mean preserving spread on the distribution of investors' endowments. Equivalently, we can model the case where the top f percent of the population has a fraction $a_R \in (0,1)$ of the wealth at the initial period by setting $\Delta(a_R) = \dfrac{1}{1-f}\left(\dfrac{a_R}{f} - 1\right)$.

Proposition 1.10. *In the subgame for given τ:*

(i) there exist Δ^* such that an equilibrium where the low-income investors follow the panic strategy do not exist for $\Delta > \Delta^*$.

(ii) there exist $\hat{f}$ and $\hat{\Delta}$ such that equilibrium where the high-income investors follow the panic strategy exists for $\Delta > \hat{\Delta}$ and $f < \hat{f}$.

According to proposition 1.10, an inequality that is sufficiently high prevents panics from spreading to the low-income investors, and at the same time, makes financial instability more prevalent for those with high-income. The reason for this outcome can be traced to the influence of the government's distributional incentives on the bailout intervention during a crisis. In particular, when inequality is sufficiently high, the haircuts imposed by the government on those with low-income - if they were to become embroiled in a bank run - are not severe enough to allow for the existence of the equilibrium where low-income investors panic. On the other hand, higher inequality would also lead to larger haircuts on high-income investors during a run on the banks, which makes them more prone to panics. Hence, a high wealth gap between the high and the low income investors would ensure the existence of an equilibrium where the former participate in a bank run. In addition, higher levels of inequality implies that the average wealth of an investor benefiting from the government bailouts is also higher.

1.4.2 Inequality and the choice of τ

The government lacks commitment and therefore the only way to influence investors' expectations about the bailout policy is through the choice of τ. For a given value of Δ, denote with $\tau_L(\Delta)$ the minimum tax rate necessary to eliminate the equilibria where the low-income investors panic and with $\tau_H(\Delta)$ – the minimum tax rate necessary to eliminate the equilibria where the high-income investors panic. That is, for $j = L, H$:

$$c_1\left(e_j(\Delta),\tau\right) < c_{2\beta}\left(e_j(\Delta),\tau\right) \quad \text{iff} \quad \tau > \tau_j(\Delta)$$

we have $\tau_L(\Delta) < \tau_H(\Delta)$ - eliminating the equilibria where the low-income investors panic requires a lower tax rate since these investors anticipate a more generous support from the government in a crisis. Given the choice of the tax rate τ, there exists an equilibrium where both the high and low income investors follow the panic strategy when $\tau \leq \tau_L(\Delta)$. On the other hand, if $\tau_L(\Delta) < \tau \leq \tau_H(\Delta)$ - an equilibrium where the low-income investors panic would no longer exists. However, the high income investors could still run on the banks in equilibrium. Finally, if $\tau > \tau_H(\Delta)$ - neither the high nor the low income investors could run on the banks in equilibrium. For given value of Δ, the government chooses the tax rate in order to maximize the sum of investors' expected utilities from the perspective of period 0. Let $\tau^*(\Delta)$ denote the tax rate chosen by the government when the wealth gap is equal to Δ.[16]

Figure 1.3 plots $\tau_L(\Delta)$, $\tau^*(\Delta)$ and $\tau_H(\Delta)$ as functions of the level of inequality in period 0.[17] When the wealth gap Δ between the high and low income investors is relatively small, the government chooses a tax rate which is enough to eliminate the equilibrium where either the high or the low income investors panic. In this situation, bank runs never occur in equilibrium and the optimal tax rate is determined solely by the desired level of the public good. As the wealth gap increases past about $\Delta = 0.1$, it is no longer the case that the optimal tax rate based solely on the desired level of the public good is high enough to prevent bank runs by wealthy agents. As the figure shows, the optimal policy is to begin increasing the tax rate τ as Δ increases in order to maintain financial stability. At a certain point, however, (about $\Delta = 0.17$ in the figure), raising the tax rate further becomes too costly and the government instead

[16]If a given tax rate is consistent with more than one equilibrium, I assume that the equilibrium that obtains involves the largest fraction of investors running on the banks. More complicated selection rules, where each of the possible equilibria is selected with a given probability, yield qualitatively similar results.

[17]The remaining parameters of the model are set as follows $R = 1.5$, $\pi = 0.5$, $\gamma = 5$, $\delta = 1$, $q = 0.05$ and $f = 0.5$.

chooses to cut the tax rate and permit runs to occur on wealthy investors' banks in state β. Figure 3 thus provides an example where more unequal wealth distribution is associated with more financial instability.

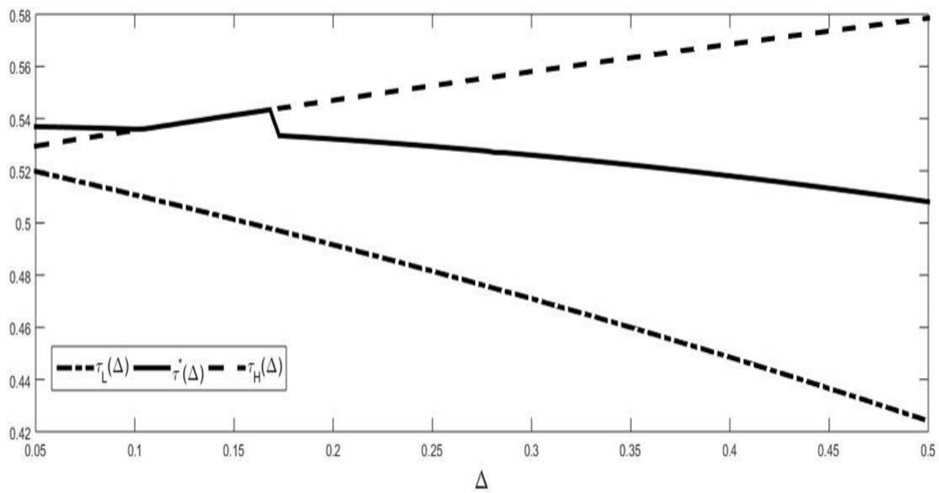


Figure 1.3: Higher inequality leading to higher fragility

In this framework, however, more unequal wealth distribution does not necessarily lead to more financial instability. In fact, we can show the following:

Proposition 1.11. *There exist economies where more unequal wealth distribution will* ***increase*** *the fragility of the financial system and economies where more unequal wealth distribution will* ***decrease*** *the fragility of the financial system.*

Table 1.1 provides an example where more unequal wealth distribution in period 0 would lower financial fragility by eliminating the equilibrium where the low-income investors panic.[18] When there is no inequality (i.e $\Delta = 0$), we have $c_1(e_L, \tau) = c_1(e_H, \tau) > c_{2\beta}(e_L, \tau) = c_{2\beta}(e_H, \tau)$ and therefore the equilibrium where all investors follow the panic strategy exists. However, when the wealth gap increases to $\Delta = 0.5$ (the initial endowment of the high-income is 1.5 times that of the low-income) the second row of the table shows that the low-income investors will not be best

[18]The parameters for the example in table 1.1 are the following $R = 3$, $\pi = 0.5$, $\gamma = 8$, $\delta = 10^{-3}$, $q = 0.05$ and $f = 0.5$.

responding with the panic strategy and therefore the equilibrium where both the high and low income investors panic in state β no longer exist. On the other hand, the third row shows that there is an equilibrium where only the high-income investors follow the panic strategy.

Inequality	Panic strategy	$c_1(e_L, \tau)$	$c_{2\beta}(e_L, \tau)$	$c_1(e_H, \tau)$	$c_{2\beta}(e_H, \tau)$	
$\Delta = 0$	high and low income	0.9838	0.8807	0.9838	0.8807	
$\Delta = 0.5$	high and low income	0.7834	0.8281	1.2136	0.9760	
$\Delta = 0.5$	high income	0.7820	0.8971	1.3056	0.9462	

Table 1.1: Higher inequality leading to lower fragility

1.4.3 Costs and benefits of more progressive taxation

Observe from Figure 1.3 that the tax rate chosen by the government is lower when inequality is higher except when taxes are kept high with the purpose of deterring high-income investors from panicking. However, if inequality becomes sufficiently high, the government would no longer find it optimal to keep taxes high in order to accomplish this objective and sets τ in the range consistent with the equilibria where high-income investors run on the banks when the state is β.

The government wants to lower the tax rate when inequality is increasing because the low-income investors will have higher marginal utility from private consumption, and therefore, imposing the same tax rate will be more costly from the perspective of the government. Therefore, by shifting the burden of taxation towards the high-income investors, a more progressive tax system in period 0 will not only have a utilitarian benefit in terms of reduced inequality, but could also prevent a decline in tax revenue and therefore preserve the capacity of the government to rescue the financial system in a crisis. However, the link between a more progressive tax system and financial fragility is subtle since we must take into account how inequality is related to financial fragility. In this framework, policies to combat the rise in inequality must incorporate the fact that more unequal wealth distribution could be associated with either an increase or a decrease in the level of financial fragility (an implication of proposition 1.11).

For the case where a more progressive tax system has an added benefit of increased financial stability consider Figure 1.3. A progressive tax system, in this case, could prevent a raise in inequality beyond the level allowing for equilibria where the high-income investors follow the panic strategy. On the other hand, a case where more progressive tax system has an added cost in terms of financial instability is shown in Table 1.1, where a highly progressive tax system imposes an additional distortion, since the equilibria where the low-income investors panic exist when inequality is at low levels.

To summarize: higher inequality combined with the government's utilitarian incentives would render financial panics easier to obtain for the high-income and harder to obtain for the low-income investors. Thus, the answer to the question of whether a period of widening inequality has the potential to increase or to decrease financial fragility hinges on whether the financial system is currently admitting equilibria where low–income investors panic and run on the banks. The answer to the previous question, at the same time, would influence the government's desired levels of tax progressivity.

1.5 Bank formation stage

The results in the previous sections were obtained under the assumption that investors with different endowments operate separate intermediation shames (type e bank for type e investors). In this section the model is augmented with a bank formation stage, which allows investors to choose where to deposit their endowment. The goal is to characterize conditions which imply that restricting attention to a separate bank for each investor type is without loss of generality.

1.5.1 Coalitions

In period 0 the government will collect taxes and then investors form coalitions. Coalitions are indexed by $k \in [0,1]$ and each investor can be part of at most one coalition. Each coalition is formed with the goal of creating an intermediation scheme (*a bank*). The bank is operating to maximize the sum of its investors' expected utilities at any point in time and must infer the realization of the sunspot state s from the withdrawal demand of its investors. In addition, the process of coalition formation is assumed to best respond, instead of trying to influence investors' withdrawal strategy. As before, I restrict attention to equilibria where investors with the same endowment follow the same withdrawal strategy. Let Ω_k denote the set of investor types e that become members of coalition k. A *coalitional structure* will be denoted by $\Omega \equiv \{\Omega_k\}_{k=0}^{1}$. In order to preserve the competitive nature of the banking sector, I restrict attention to *competitive* coalitional structures where no single coalition has the capacity to influence economy-wide outcomes and I assume that all coalitions that can potentially form are competitive. The payment schedule for type e investor in coalition k is denoted $C^k(e,\tau)$:

$$C_k(e,\tau) \equiv \left(c_1^k(e,\tau), c_{2\alpha}^k(e,\tau), c_{1\beta}^k(e,\tau), c_{2\beta}^k(e,\tau)\right) \tag{1.38}$$

and the collection of the payments schedules for all investor types in coalition k will be denoted $C_k(\tau) \equiv \{C_k(e,\tau)\}_{e\in\Omega_k}$. The expected utility from private consumption for a type e investor from becoming a member of coalition k is thus:

$$W_k(e,\tau) = \left\{ \begin{array}{c} \pi u\left(c_1^k(e,\tau)\right) + (1-q)(1-\pi)u\left(c_{2\alpha}^k(e,\tau)\right) \\ +q(1-\pi)\left(\pi_\beta(e)u\left(c_{1\beta}^k(e,\tau)\right) + (1-\pi_\beta(e))u\left(c_{2\beta}^k(e,\tau)\right)\right) \end{array} \right\} \tag{1.39}$$

where $\pi_\beta(e) = 0$ if type e investors follow the no-panic withdrawal strategy and $\pi_\beta(e) = \pi$ if type e investors follow the panic withdrawal strategy. In state β, coalitions that among their members have investors following the panic strategy will be experiencing a run and will be eligible for a bailout transfer from the government. Let $b_k(\tau) \geq 0$ denote the per-capita bailout to coalition k. The payments to the remaining investors in coalitions with a bailout will be mandated by the government in order to ensure the ex-post efficient allocation of consumption across the private and the public good:

$$c^k_{2\beta}(e,\tau) = min\left\{c^G_{2\beta}(\tau), c^k_{2\alpha}(e,\tau)\right\} \tag{1.40}$$

$$u'(c^k_{1\beta}(e,\tau)) = Ru'(c^k_{2\beta}(e,\tau)) \tag{1.41}$$

$$u'(c^G_{1\beta}(\tau)) = v'\left(\tau - \overset{[}{0}]1 \int b_k(\tau)dk\right) \tag{1.42}$$

The government mandated payments are restricted not to exceed the payments to patient investors in case their coalition was not experiencing a run, i.e. $c^k_{2\beta}(e,\tau) \leq c^k_{2\alpha}(e,\tau)$ for $e \in \Omega_k$. That is, the possibility where some investors benefit from a run on the banks is not allowed. At the same time, the remaining components of the payment schedule in each coalition are restricted to satisfy a *proportional rule*: for $\{e_1, e_2\} \subseteq \Omega_k$:

$$\frac{c^k_1(e_1,\tau)}{c^k_1(e_2,\tau)} = \frac{c^k_{2\alpha}(e_1,\tau)}{c^k_{2\alpha}(e_2,\tau)} = \frac{e_1}{e_2} \tag{1.43}$$

In addition, if the coalition is not bailed out, the payment schedule in state β must also obey this rule:

$$\frac{c^k_{1\beta}(e_1,\tau)}{c^k_{1\beta}(e_2,\tau)} = \frac{c^k_{2\beta}(e_1,\tau)}{c^k_{2\beta}(e_2,\tau)} = \frac{e_1}{e_2} \quad if \quad b_k(\tau) = 0 \tag{1.44}$$

Observe that this proportional rule will hold trivially when all investors in a given coalition have the same endowment. In addition, (1.43) implies that all patient investors within the coalition receive the same rate of return on their deposit if their coalition does not experience a run.

Definition 1.12. A coalition k follows the *proportional rule* if for $\{e_1, e_2\} \subseteq \Omega_k$ the payment schedules for type e_1 and type e_2 investors - $C_k(e_1, \tau)$, $C_k(e_2, \tau)$ - satisfy:

(i) (1.43) - (1.44) when $b_k(\tau) = 0$

(ii) (1.40) - (1.43) when $b_k(\tau) > 0$

Henceforth, I restrict attention to coalitions which follow the proportional rule.

1.5.2 Endogenous segmentation of the banking sector

A coalitional structure Ω is *stable* (alternatively belongs to the *core*), if no coalition within this structure is *blocked*. A given coalition $\Omega_k \in \Omega$ is *blocked* whenever a new coalition can be formed that would make a fraction of the investors in k strictly better of in terms of expected utility in period 0.[19] Next, I denote with

$$C_S(e,\tau) \equiv \left(c_1^S(e,\tau), c_{2\alpha}^S(e,\tau), c_{1\beta}^S(e,\tau), c_{2\beta}^S(e,\tau)\right) \tag{1.45}$$

the payment schedule for type e investors when they operate an intermediation scheme on their own and with $W_S(e,\tau)$ - the expected utility from private consumption in period 0 associated with the payment schedule in (1.45). As in section 1.3, $C_S(e,\tau)$ will be characterize by (1.10) - (1.11), (1.14) - (1.15) and (1.22) if the government intervenes and provides a bailout in state β and by (1.10) - (1.11) and (1.23) otherwise. Note that a necessary condition for a given coalitional structure Ω to be stable is that for each $\Omega_k \in \Omega$ and for each $e \in \Omega_k$ we have:

$$W_k(e,\tau) \geq W_S(e,\tau)$$

[19] This blocking criteria (also called objections criteria) is widely used in the literature on coalition formation to reduce the set of admissible outcomes of the coalition formation process (Ray and Vohra 2015 provide a survey of the relevant literature). In this setting, a stable coalition structure also satisfies the farsighted refinement of the blocking criteria introduced by Harsanyi (1974).

Next, let $\Omega^* \equiv \{\Omega_k^*\}_{k=0}^1$ be a stable coalition structure and let $\{C_k^*(\tau)\}_{k=0}^1$ be the collection of payment schedules associated with Ω^* (where $C_k^*(\tau) = \{C_k^*(e,\tau)\}_{e\in\Omega_k}$). Denote with $\{C_S(e,\tau)\}_{e\in[e_L,e_H]}$ the collection of payment schedules associated with a separate coalition for each investor type. I will refer to the coalition structure Ω as being *payoff-equivalent to a separate coalition for each investor type* if for each $e \in [e_L, e_H]$ and $C_k(e,\tau) \in \{C_k(\tau)\}_{k=0}^1$ we have $C_k(e,\tau) = C_S(e,\tau)$. We are now ready to state the main result of this section.

Proposition 1.13. ***(Payoff-equivalence)*** *If Ω^* is a stable coalition structure, then Ω^* is payoff-equivalent to a separate coalition for each investor type*

In other words, restricting attention to type-specific banks, as was done in Sections 1.3 and 1.4, is without loss of generality. The intuition for Proposition 1.13 is the following: a given coalition involving investors that follow different withdrawal strategies will be blocked by a coalition that involves only those investors following the no-panic strategy. At the same time, any coalition involving investors who follow the same withdrawal strategy (either the panic or the no-panic) behaves as if each investor type within the coalition was operating their own liquidity insurance arrangement (a type e bank for type e investors).

Before concluding this section, observe that, even though the proportional rule might be a reasonable way to model bank's behavior, especially in environments with limited commitment where more complicated rules could fail to be credible, we can formulate a number of other mechanisms designed to allocate consumption to the investors. Nonetheless, in order to preserve the main massage of the paper, we do not require strict banking segmentation, but rather, two conditions, both of which are empirically relevant: *(i)* a financial system where investors' claim on intermediaries is positively correlated with their wealth, and *(ii)* a government that is able to channel funds to specific investors based on their observable characteristics (i.e. wealth).

1.6 Conclusion

I have presented a model of financial intermediation in which the government is unable to commit to the details of the bailout intervention before the the crisis and where investors have different wealth levels. I showed that an equilibrium where wealthy investors run on the banks and prompt the government to bail them out can arise naturally in setting such as these. The reason for this outcome is that the rescue program of the government will be motivated, at least in part, by distributional incentives and will induce endogenous caps to deposit insurance. These caps, in turn, will be decreasing as a function of investors' wealth levels. Thus, the model predicts that financial institutions with wealthy investors will be more likely to benefit from bailouts - an outcome occurring in the absence of any political rent-seeking frictions and under a government that is both benevolent and utilitarian.

If inequality is higher, then the government's distributional incentives during the implementation of the bailout program will also become more pronounced which, in turn, increases wealthy investors' incentive to panic. In settings such as these, a more progressive tax code could lower the level of inequality and prevent equilibria where wealthy investors end up engulfed in a banking panic. Furthermore, a government that is able to use a more progressive tax code in order to shift the burden of taxation towards the wealthy will be less eager to cut taxes when inequality is higher. Hence, by preventing a decline in the government's tax revenue and preserving the government's capacity to rescue the financial system in a crisis, a more progressive tax code could have a prudential effect on financial stability. However, the same distributional incentives of the government imply that effect of inequality on financial fragility is not straightforward. In some cases, higher inequality could actually increase financial stability, since those parts of the financial system providing intermediation to the relatively poor would benefit from a more generous government support in a crisis and might no longer be susceptible to self-fulfilling runs.

The framework in this paper also accounted for the moral hazard issues stemming from the bailout intervention in a crisis. In particular, banks that anticipate to be bailed out ignore the cost that will be imposed on the public sector in a financial panic and as a result engage in excessive levels of maturity transformation. At the same time, a novel aspect of the approach followed here is that a run on a given bank will constitutes a necessary but not a sufficient condition for a government intervention. In fact, a given bank would anticipate the conditions leading the government to bail it it out and would decide ex-ante whether to qualify for a bailout in case of a run from its investors. That is, whether or not a financial panic also leads to government intervention is an endogenous outcome of the model.

I conclude by outlining three potentially promising directions for future research. *First*, one of the tacit assumptions in the paper is that the financial system is relatively transparent and allows the government to determine investors' wealth before the implementation of the bailout policy. However, given the main massage of the paper - namely that the most sizable haircuts in a financial crisis would tend to be imposed on the wealthy - those investors would also be eager to find ways to increase the opacity of the financial system and to make it harder for the government to implement its desired bailout intervention. *Second*, preventing contagion is often given as one of the primary reasons for the government's rescue program. Extending the model to account for this additional dimension implies that the government must balance its distributional concerns with its desire to prevent contagion. *Third*, in this model, appointing a bailout authority that is more inclined to rescue financial institutions with wealthy investors (rather than being strictly utilitarian) could, in some cases, eliminate the equilibria where they panic and require subsequent government support – an outcome which would benefit all agents. This implication is, in fact, similar to the literature on the optimal monetary policy where appointing a central banker whose commitment to fight inflation is stronger than that of the public could be beneficial and increase equilibrium welfare (Rogoff 1985). However, such a policy could introduce other distortions to the extent that this person or agency has the ability to influence policy decisions in normal times. Exploring these dimensions of the model is an interesting topic for future research.

1.7 Appendix I: Proofs

Proposition 1.2.

Proof. First, suppose that both type e_1 and type e_2 banks best respond by qualifying for a bailout. The payment schedule for type e_1 and type e_2 investors in this case will be determined by (1.10), (1.11) and (1.22). Whereas the bailout to type e_1 and type e_2 banks by (1.16). In order to establish the desired result, it is sufficient to show:

$$c_1(e_1, \tau) < c_1(e_2, \tau) \quad for \quad e_1 < e_2$$

since,(1.16) will then imply that $b(e_1, \tau) > b(e_2, \tau)$. The proof is by contradiction. Assume it were the case that $c_1(e_1, \tau) \geq c_1(e_2, \tau)$, then since $e_1 < e_2$ we have:

$$(1 - \tau)e_1 - \pi c_1(e_1, \tau) < (1 - \tau)e_1 - \pi c_1(e_2, \tau)$$

Given that both type e_1 and type e_2 follow the panic strategy, (1.6) implies that $\pi_\alpha(e_1) = \pi_\alpha(e_2)$ and $\pi_\beta(e_1) = \pi_\beta(e_2)$ and we obtain from the budget constraint in state α:

$$c_{2\alpha}(e_1, \tau) < c_{2\alpha}(e_2, \tau)$$

The first order condition in (1.11) implies:

$$\mu_\alpha(e_1, \tau) > \mu_\alpha(e_2, \tau)$$

Then from (1.22) we obtain $c_1(e_1,\tau) < c_1(e_2,\tau)$ - a contradiction since we initially assumed that $c_1(e_1,\tau) \geq c_1(e_2,\tau)$. Therefore we must have:

$$c_1(e_1,\tau) < c_1(e_2,\tau)$$

and we obtain the desired result from (1.16), namely $e_1 < e_2$ implies $b(e_1,\tau) > b(e_2,\tau)$.

Second, Proposition 1.5 implies that type e_1 banks would best respond by qualifying for a bailout whenever type e_2 best respond by qualifying for a bailout. Therefore, either only type e_1 banks qualify for a bailout, in which case $b(e_1,\tau) > 0 = b(e_2,\tau)$. Or both banks qualify for a bailout, in which case we have established $b(e_1,\tau) > b(e_2,\tau)$. □

Proposition 1.3.

Proof. Denote with $C_B(e,\tau)$ the set of all possible payments during the first π withdrawals for which a type e bank will qualify for a bailout when experiencing a run.

$$C_B(e,\tau) = \left\{c_1(e,\tau) \ \ s.t. \ (1-\tau)e - \pi c_1(e,\tau) < \psi_\beta^B(\tau)\right\}$$

The optimal choice of $c_1(e,\tau)$ when qualifying for a bailout solves the following program:

$$W^B(e,\tau) = \underset{c_1(e,\tau)\in C_B(e,\tau)}{max} \left\{ \begin{array}{c} \pi u(c_1(e,\tau)) + (1-q)V((1-\tau)e - \pi c_1(e,\tau);\, \pi_\alpha) \\ +qV\left(\psi_\beta^B(\tau);\, \pi_\beta\right) \end{array} \right\}$$

Next, denote with $C_{NB}(e,\tau)$ the set of all possible payments for which a type e bank will not qualify for bailout when experiencing a run (i.e. it self-insures):

$$C_{NB}(e,\tau) = \left\{c_1(e,\tau) \ \ s.t. \ (1-\tau)e - \pi c_1(e,\tau) \geq \psi_\beta^B(\tau)\right\}$$

The optimal choice of $c_1(e,\tau)$ in this case solves the following program:

$$W^{NB}\left(e,\tau\right)=\underset{c_1(e,\tau)\in C_{NB}(e,\tau)}{argmax}\left\{\begin{array}{c}\pi u\left(c_1\left(e,\tau\right)\right)+(1-q)V\left((1-\tau)e-\pi c_1\left(e,\tau\right);\,\pi_\alpha\right)\\ +qV\left((1-\tau)e-\pi c_1(e,\tau);\,\pi_\beta\right)\end{array}\right\}$$

If $c_1^B(e,\tau)$ is characterized by (1.22) and if $c_1^{NB}(e,\tau)$ is characterized by (1.23), then Proposition 1.5 implies $c_1^{NB}(e,\tau) < c_1^B(e,\tau)$ and, in order to establish the desired result, we must consider three cases.

Case 1: $c_1^B(e,\tau)$ triggers a bailout, whereas $c_1^{NB}(e,\tau)$ would not trigger a bailout:

$$(1-\tau)e-\pi c_1^B(e,\tau)<\psi_\beta^B(\tau)<(1-\tau)e-\pi c_1^{NB}(e,\tau)$$

Then, a type e bank best responding by qualifying for a bailout sets $c_1(e,\tau)$ as in (1.22), given that this is the optimal choice of $c_1(e,\tau)$ without the additional constraint $c_1(e,\tau)\in C_B\left(e,\tau\right)$. Similarly, a type e bank best responding by self-insuring would set $c_1(e,\tau)$ as in (1.23) given that this is the optimal choice of $c_1(e,\tau)$ without imposing the constraint $c_1\left(e,\tau\right)\in C_{NB}\left(e,\tau\right)$.

Case 2: Both $c_1^B(e,\tau)$ and $c_1^{NB}(e,\tau)$ would trigger a bailout:

$$(1-\tau)e-\pi c_1^B(e,\tau)<(1-\tau)e-\pi c_1^{NB}(e,\tau)<\psi_\beta^B(\tau)$$

If a type e bank is self-insuring, then $\hat{c}_1^{NB}(e,\tau)$ must be set to satisfy:

$$(1-\tau)e-\pi\hat{c}_1^{NB}(e,\tau)\geq\psi_\beta^B(\tau)$$

hence $\hat{c}_1^{NB}(e,\tau)<c_1^{NB}(e,\tau)$. I show that in this case type e banks strictly prefer to qualify for a bailout. Indeed, consider the following:

$$\pi u\left(c_1^B(e,\tau)\right) + (1-q)V\left((1-\tau)e - \pi c_1^B(e,\tau);\, \pi_\alpha\right) + qV\left(\psi_\beta^B(\tau);\, \pi_\beta\right)$$
$$>$$
$$\pi u\left(c_1^{NB}(e,\tau)\right) + (1-q)V\left((1-\tau)e - \pi c_1^{NB}(e,\tau);\, \pi_\alpha\right) + qV\left(\psi_\beta^B(\tau);\, \pi_\beta\right)$$
$$\geq$$
$$\pi u\left(c_1^{NB}(e,\tau)\right) + (1-q)V\left((1-\tau)e - \pi c_1^{NB}(e,\tau);\, \pi_\alpha\right) + qV\left((1-\tau)e - \pi c_1^{NB}(e,\tau);\, \pi_\beta\right)$$
$$>$$
$$\pi u\left(\hat{c}_1^{NB}(e,\tau)\right) + (1-q)V\left((1-\tau)e - \pi \hat{c}_1^{NB}(e,\tau);\, \pi_\alpha\right) + qV\left((1-\tau)e - \pi \hat{c}_1^{NB}(e,\tau);\, \pi_\beta\right)$$

The first inequality follows from the fact that $c_1^B(e,\tau)$ is the optimal unconstrained choice of $c_1(e,\tau)$ when the bank is bailed out (and therefore $\psi_\beta^B(\tau)$ is treated as exogenous by the bank). The second inequality follows from the assumption that setting $c_1^{NB}(e,\tau)$ as in (1.23) leads to $(1-\tau)e - \pi c_1^{NB}(e,\tau) < \psi_\beta^B(\tau)$. The third inequality from the fact that $c_1^{NB}(e,\tau)$ is the optimal choice of early payments when the bank is not bailed out and there are no additional restrictions on the choice of $c_1(e,\tau)$. Therefore, type e banks would best respond by qualifying for a bailout and their choice of $c_1(e,\tau)$ will be characterized by (1.22).

Case 3: Both $c_1^B(e,\tau)$ and $c_1^{NB}(e,\tau)$ would not trigger a bailout

$$(1-\tau)e - \pi c_1^{NB}(e,\tau) \geq (1-\tau)e - \pi c_1^B(e,\tau) \geq \psi_\beta^B(\tau)$$

Therefore, in order for a type e bank to qualify for a bailout, it needs to set $\hat{c}_1^B(e,\tau)$ such that:

$$(1-\tau)e - \pi \hat{c}_1^B(e,\tau) < \psi_\beta^B(\tau)$$

and we have $\hat{c}_1^B(e,\tau) > c_1^B(e,\tau)$. But this would also imply that type e banks are strictly better off by self-insuring. Indeed, consider the following:

$$\pi u\left(c_1^{NB}(e,\tau)\right)+(1-q)V\left((1-\tau)e-\pi c_1^{NB}(e,\tau);\,\pi_\alpha\right)+qV\left((1-\tau)e-\pi c_1^{NB}(e,\tau);\,\pi_\beta\right)$$
$$>$$
$$\pi u\left(c_1^{B}(e,\tau)\right)+(1-q)V\left((1-\tau)e-\pi c_1^{B}(e,\tau);\,\pi_\alpha\right)+qV\left((1-\tau)e-\pi c_1^{B}(e,\tau);\,\pi_\beta\right)$$
$$\geq$$
$$\pi u\left(c_1^{B}(e,\tau)\right)+(1-q)V\left((1-\tau)e-\pi c_1^{B}(e,\tau);\,\pi_\alpha\right)+qV\left(\psi_\beta^B(\tau);\,\pi_\beta\right)$$
$$>$$
$$\pi u\left(\hat{c}_1^{B}(e,\tau)\right)+(1-q)V\left((1-\tau)e-\pi \hat{c}_1^{B}(e,\tau);\,\pi_\alpha\right)+qV\left(\psi_\beta^B(\tau);\,\pi_\beta\right)$$

Neither $c^{NB}(e,\tau)$ nor $c^{B}(e,\tau)$ leads to a bailout in state β. However, $c^{NB}(e,\tau)$ is also the unconstrained optimal choice of $c(e,\tau)$ when type e banks self-insure and therefore we obtain the first inequality. The second inequality follows from $(1-\tau)e-\pi c_1^B(e,\tau)\geq\psi_\beta^B(\tau)$. Whereas the third inequality follows from the fact that $c^B(e,\tau)$ is the unconstrained optimal choice of early payments when $\psi_\beta^B(\tau)$ is viewed as exogenous by the bank. Therefore, type e banks best respond by choosing to self-insure and will set $c_1(e,\tau)$ as in (1.23). □

The discussion in section 1.3.5, in particular, conditions (1.24) and (1.25) rely on the following lemma.

Lemma 1.

Proof. Suppose that type e banks best respond by qualifying for a bailout in state β, then $c_1(e,\tau)$ is characterized by (1.23), combined with (1.11) this yields:

$$u'(c_1^B(e,\tau))=(1-q)Ru'(c_{2\alpha}^B(e,\tau))$$

A necessary condition for the panic strategy in (1.2) to be a best respond is $c_{2\alpha}^B(e,\tau)>c_1^B(e,\tau)$, which is the case whenever:

$$q<\frac{R-1}{R}$$

but this is the same as condition (1.24). On the other hand, suppose that type e banks best respond by self-insuring. By combining (1.23) with (1.11) we obtain:

$$u'(c_1^{NB}(e,\tau)) = (1-q)Ru'(c_{2\alpha}^{NB}(e,\tau)) + qRu'(c_{2\beta}^{NB}(e,\tau))$$

Next, if type e investors best respond with the panic strategy, we must have $c_1^{NB}(e,\tau) > c_{2\beta}^{NB}(e,\tau)$. From from (1.23) and (1.11), this is equivalent to:

$$(1-q)Ru'(c_{2\alpha}^{NB}(e,\tau)) + qRu'(c_{2\beta}^{NB}(e,\tau)) < u'(c_{2\beta}^{NB}(e,\tau))$$

or equivalently:

$$\frac{u'(c_{2\alpha}^{NB}(e,\tau))}{u'(c_{2\beta}^{NB}(e,\tau))} < \frac{1-Rq}{R-Rq}$$

Given that investors utility is constant relative risk aversion with parameter $\gamma > 1$, the above holds if and only if the parameters of the model satisfy:

$$\frac{R^{\frac{1-\gamma}{\gamma}}}{\pi + (1-\pi)R^{\frac{1-\gamma}{\gamma}}} < \left(\frac{1-Rq}{R-Rq}\right)^{\frac{1}{\gamma}}$$

which is the same as condition (1.25). $\square$

Proposition 1.4.

Proof. We must show that $c_1^{B}(e,\tau) > c_1^{NB}(e,\tau)$. The proof is by contradiction. Suppose that

$$c_1^{B}(e,\tau) \leq c_1^{NB}(e,\tau)$$

then

$$u'(c_1^{B}(e,\tau)) \geq u'(c_1^{NB}(e,\tau))$$

From Proposition 1.3, $c_1^B(e,\tau)$ is characterized by

$$u'(c_1^B(e,\tau)) = (1-q)\mu_\alpha^B(e,\tau)$$

whenever a bank best responds by qualifying for a bailout and by

$$u'(c_1^{NB}(e,\tau)) = (1-q)\mu_\alpha^{NB}(e,\tau) + q\mu_\beta^{NB}(e,\tau)$$

whenever a bank best responds by self-insuring. From (1.22) and (1.23) we obtain:

$$(1-q)\mu_\alpha^B(e,\tau) \geq (1-q)\mu_\alpha^{NB}(e,\tau) + q\mu_\beta^{NB}(e,\tau)$$

and since $\mu_\beta^{NB}(e,\tau) > 0$, we must have:

$$\mu_\alpha^B(e,\tau) > \mu_\alpha^{NB}(e,\tau)$$

Combining the above expression with (1.11) yields:

$$c_{2\alpha}^B(e,\tau) < c_{2\alpha}^{NB}(e,\tau)$$

Working with the budget constraint in state α:

$$\begin{aligned}(1-\tau)e &= \pi c_1^B(e,\tau) + (1-\pi)\frac{c_{2\alpha}^B(e,\tau)}{R} \\ &< \pi c_1^{NB}(e,\tau) + (1-\pi)\frac{c_{2\alpha}^{NB}(e,\tau)}{R} \\ &= (1-\tau)e\end{aligned}$$

i.e. $(1-\tau)e < (1-\tau)e$ which is a contradiction. Therefore the initial assumption $c_1^B(e,\tau) \leq c_1^{NB}(e,\tau)$ cannot be true and we must have $c_1^B(e,\tau) > c_1^{NB}(e,\tau)$ as desired. □

Proposition 1.5.

Proof. Follows from the discussion in the text. □

Proposition 1.6.

Proof. Suppose that $X(\tau) \neq \emptyset$, let $x \in X(\tau)$, and consider a panic set of investor types which is given by $P(\tau) = [x, e_H]$. That is, investors with endowment $e_L \leq e < x$ follow the no panic strategy in (1.1) whereas investors with endowment $x \leq e \leq e_H$ follow the panic strategy in (1.2).

First, if type e investors follow the no-panic strategy, then from (1.5) only impatient investors are among those to withdraw before the state is inferred by type e banks, hence $\pi_\alpha(e) = \pi_\beta(e) = 0$. In addition, type e banks do not experience a run and therefore do not receive a bailout. From (1.7), (1.8), (1.11) and (1.23) we obtain that the payment schedule in banks whose investors do not panic is not contingent on s:

$$c_1^{NP}(e,\tau) = c_{1\beta}^{NP}(e,\tau)$$

$$c_2^{NP}(e,\tau) \equiv c_{2\beta}^{NP}(e,\tau) = c_{2\alpha}^{NP}(e,\tau)$$

where $c_1^{NP}(e,\tau)$ and $c_2^{NP}(e,\tau)$ will be the solution to:

$$\pi c_1^{NP}(e,\tau) + (1-\pi)\frac{c_2^{NP}(e,\tau)}{R} = (1-\tau)e$$

$$u'\left(c_1^{NP}(e,\tau)\right) = Ru'\left(c_2^{NP}(e,\tau)\right)$$

Observe that these two equations are standard in a Diamond and Dybvig model. Moreover, since $R > 1$ we have $c_1^{NP}(e,\tau) < c_2^{NP}(e,\tau)$ i.e. all types that follow the no-panic strategy will be best responding.

Second, consider type x investors. Since $x \in X(\tau)$ we have from (1.28):

$$D\left(x, \tau; [x, e_H]\right) > 0$$

i.e. type x banks best respond by qualifying for a bailout. In addition, from (27), there must exist e^{NB} such that:

$$D\left(e^{NB},\tau;[x,e_H]\right)\geq 0 \quad \text{with} \quad ''='' \quad \text{if} \quad e^{NB}<e_H$$

Further, from (1.27) it follows that the above function is decreasing in e. Thus, we have $D\left(e^{NB},\tau;[x,e_H]\right) > 0$ for $x \leq e < e^{NB}$ and $D\left(e^{NB},\tau;[x,e_H]\right) < 0$ for $e^{NB} < e \leq e_H$. Therefore, banks servicing investors with $x \leq e < e^{NB}$ qualify for a bailout and set $c_1(e,\tau)$ as in (1.22). On the other hand, banks servicing investors with $e^{NB} < e \leq e_H$ prefer to self-insure and set $c_1(e,\tau)$ as in (1.23). Next, we must show that all investors with endowment in the interval $[x,e_H]$ are best responding with the panic strategy in (1.2). Since $x \in X(\tau)$ we obtain from (1.28):

$$c_1^B(x,\tau) \geq c_{2\beta}^B(\tau;[x,e_H])$$

that is, type x investors have an incentive to panic in state β. From (1.10), (1.11) and (1.22) we obtain that $c_1^B(e,\tau)$ is an increasing function of e and hence, for each e such that $x < e < e^{NB}$:

$$c_1^B(e,\tau) > c_{2\beta}^B(\tau;[x,e_H])$$

hence all investors with endowment in the interval $\left(x,e^{NB}\right)$ also have an incentive to panic in state β. In addition, (1.24) implies that:

$$c_1^B(e,\tau) < c_{2\alpha}^B(e,\tau)$$

therefore all investors whose endowment belongs to the interval $e \in [x,e^{NB})$ will be best responding with the panic strategy in (1.2). Finally, from (1.25) we obtain that the condition in (1.4) is satisfied in all banks choosing to self-insure and therefore investors whose endowment exceeds e^{NB} will also be best responding with the panic strategy. $\square$

Propositions 1.7 and 1.8 use the following lemma.

Lemma 2. If $P(\tau) \subset P^*(\tau)$, then $b(e,\tau;P(\tau)) \geq b(e,\tau;P^*(\tau))$ for $e \in P(\tau) \cap P^*(\tau)$

Proof. The proof is by contradiction: suppose this were not true. That is, there exist $\hat{e} \in P(\tau)$ such that:

$$b\left(\hat{e},\tau,P(\tau)\right) < b\left(\hat{e},\tau,P^*(\tau)\right) \tag{1.46}$$

From, (1.10), (1.11) and (1.22) it follows that $c_1^B(\hat{e},\tau)$ - the payment to type $\hat{e}$ investor withdrawing before the state is inferred by the banks - is the same, when the panic set is either $P(\tau)$ or $P^*(\tau)$. Hence, we have:

$$(1-\tau)\,e - \pi c_1^B(\hat{e},\tau) + b\left(\hat{e},\tau;P(\tau)\right) < (1-\tau)\,e - \pi c_1^B(\hat{e},\tau) + b\left(\hat{e},\tau;P^*(\tau)\right)$$

and from (1.14) and (1.15) we obtain:

$$\psi_\beta^B\left(\tau,P(\tau)\right) < \psi_\beta^B\left(\tau;P^*(\tau)\right)$$

from (1.10) and (1.11) it follows that:

$$c_{1\beta}^B\left(\tau,P(\tau)\right) < c_{1\beta}^B\left(\tau,P^*(\tau)\right) \tag{1.47}$$

By applying (1.16) the existence of such a $\hat{e}$ would imply that for all $e \in P(\tau)$:

$$b\left(e,\tau\,;P(\tau)\right) < b\left(e,\tau\,;P^*(\tau)\right)$$

and this would imply that the aggregate bailout when the panic set of investor type is $P(\tau)$ is lower than to the aggregate bailout when the panic set of investor type is $P^*(\tau)$. Indeed, we have:

$$\begin{aligned} B(\tau\,;P(\tau)) &= \int_{e\in P(\tau)} b(e,\tau,P(\tau))dG \\ &< \int_{e\in P(\tau)} b(e,\tau,P^*(\tau))dG \\ &\leq \int_{e\in P^*(\tau)} b^*(e,\tau,P^*(\tau))dG \\ &= B^*(\tau\,;P^*(\tau)) \end{aligned} \tag{1.48}$$

where the first inequality follows from $b(e,\tau,P(\tau)) < b(e,\tau,P^*(\tau))$ and the second inequality from our assumption that $P(\tau) \subset P^*(\tau)$. For fixed τ, we have:

$$\tau - B(\tau\,;P(\tau)) > \tau - B(\tau\,;P^*(\tau))$$

and then (1.11) and (1.14) yields:

$$c^B_{1\beta}\left(\tau, P(\tau)\right) > c^B_{1\beta}\left(\tau, P^*(\tau)\right) \tag{1.49}$$

which is a contradiction because (1.47) and (1.49) cannot both be true. Thus, our initial assumption, namely that there exist an investor type $\hat{e}$ such that $\hat{e} \in P(\tau)$ and whose bank receive larger per-capita bailout when the panic set is $P^*(\tau)$ cannot be true and we conclude that for all $e \in P(\tau)$ we must have:

$$b\left(e,\tau\,;P(\tau)\right) \geq b\left(e,\tau\,;P^*(\tau)\right)$$

as desired. □

Proposition 1.7.

Proof. $D\left(y,\tau;[y,e_H]\right) \leq 0$ implies that for $e > y$ we have $D\left(y,\tau;[y,e_H]\right) < 0$ and therefore banks will best respond by self-insuring. At the same time, the condition in (1.25) implies that investors in banks that are not bailed out best respond with the panic strategy. Finally, given that all banks best respond by self-insuring, there will be no bailouts in equilibrium.

Next, I will derive some properties of $X(\tau)$ and $Y(\tau)$

(i) For each $\tau \in [0,1)$ we have $X(\tau) \cap Y(\tau) = \emptyset$.

To see this, suppose there exist $\tau^* \in [0,1)$ s.t. $X(\tau^*) \cap Y(\tau^*) \neq \emptyset$ and let $z \in X(\tau^*) \cap Y(\tau^*)$, then we must have $D(z, \tau^*; [z, e_H]) > 0$ and, at the same time, $D(z, \tau^*; [z, e_H]) \leq 0$ - which is not possible and therefore $X(\tau) \cap Y(\tau) = \emptyset$.

(ii) If $x \in X(\tau)$ then each equilibrium panic set $P(\tau)$ such that $P(\tau) \subseteq [x, e_H]$ must be associated with bailouts in state β.

From (i) we have $x \in X(\tau) \Rightarrow x \notin Y(\tau)$ and therefore the sub-game for τ and a panic set $[x, e_H]$ would involve bailouts when the state is β. Next, applying Lemma 2 for $e \in P(\tau) \subset [x, e_H]$, we have $b(e, \tau; [x, e_H]) \leq b(e, \tau; P(\tau))$, which implies $c_{1\beta}^B(\tau; [x, e_H]) \leq c_{1\beta}^B(\tau; P(\tau))$ and therefore

$$0 < D(e, \tau; [x, e_H]) \leq D(e, \tau; P(\tau))$$

That is, if type e banks best respond by qualifying for a bailout when the panic set is $[x, e_H]$ they would also best respond by qualifying for a bailout when the panic set is $P(\tau) \subset [x, e_H]$.

(iii) If $X(\tau) = \emptyset$ and $Y(\tau) = \emptyset$ then a panic equilibrium does not exist for the given τ.

Suppose there is $P(\tau) \neq \emptyset$ consistent with equilibrium and let

$$z \equiv \{e \in P(\tau)\ s.t\ e \leq e'\ for\ each\ e' \in P(\tau)\}$$

we must have $P(\tau) \subseteq [z, e_H]$. From (ii), if type z banks qualify for a bailout when the panic set is $[z, e_H]$, they would also qualify for a bailout when the panic set is $P(\tau)$. In addition, from (ii) we have $c_{1\beta}^B(\tau; [z, e_H]) \leq c_{1\beta}^B(\tau; P(\tau))$. Finally, since $X(\tau) = \emptyset$, we must have $c_1^B(z, \tau) < c_{2\beta}^B(\tau; [z, e_H])$ and therefore $c_1^B(z, \tau) < c_{2\beta}^B(\tau; P(\tau))$, which implies that $P(\tau)$ cannot be an equilibrium panic set because type z investors are not best responding with the panic strategy. $\square$

Proposition 1.8.

Proof. I will show that any panic set $P(\tau) \subseteq [e_L, e_H]$ such that $e \in P(\tau)$ and $e_L \leq e < e^T(\tau)$ will not be consistent with equilibrium for the given value of τ. The proof is divided into two steps.

Step 1 Show that for any e^* such that $e_L \leq e^* < e^T(\tau)$ the panic set $P(\tau) = [e^*, e_H]$ cannot be part of equilibrium. We must show that there exist $e \in [e^*, e_H]$ such that:

$$c_1^B(e, \tau) < c_{2\beta}^B(\tau;\, [e^*, e_H])$$

which would imply that type e investors are not best responding with the panic strategy. The proof is by contradiction: suppose there exist $e_1 \in [e_L, e^T(\tau))$ such that the panic set $[e_1, e_H]$ is consistent with equilibrium in the sub-game for τ. We have must have:

$$c_1^B(e_1, \tau) \geq c_{2\beta}^B(\tau;\, [e_1, e_H])$$

Next, we must have $e_1 > e_L$. The reason is that by assumption

$$c_1^B(e_L, \tau) < c_{2\beta}^B(\tau;\, [e_L, e_H])$$

which implies that $[e_1, e_H]$ is not an equilibrium panic set when $e_1 = e_L$. Also, since $e_1 < e^T(\tau)$ it follows that e_1 is not a threshold endowment type and therefore:

$$c_1^B(e_1, \tau) > c_{2\beta}^B(\tau;\, [e_1, e_H])$$

Since $e_1 > e_L$ we have $[e_1, e_H] \subset [e_L, e_H]$ and by lemma 2:

$$c_{2\beta}^B(\tau;\, [e_L, e_H]) \leq c_{2\beta}^B(\tau;\, [e_1, e_H])$$

which implies:

$$c_1^B(e_L, \tau) < c_{2\beta}^B(\tau;\, [e_1, e_H])$$

and we have

$$c_1^B(e_L, \tau) < c_{2\beta}^B(\tau;\, [e_1, e_H]) \quad \text{and} \quad c_1^B(e_1, \tau) > c_{2\beta}^B(\tau;\, [e_1, e_H])$$

Next, since $c_1^B(e, \tau)$ is an increasing function of e there must exist e_2 s.t. $e_L < e_2 < e_1$ and

$$c_1^B(e_2, \tau) = c_{2\beta}^B(\tau;\, [e_1, e_H])$$

and we have:

$$c_1^B(e_2, \tau) = c_{2\beta}^B(\tau;\, [e_1, e_H]) \geq c_{2\beta}^B(\tau;\, [e_2, e_H])$$

where the last inequality follows from $e_2 < e_1$ and Lemma 2. Next, e_2 is not a threshold endowment since $e_2 < e^T(\tau)$ and hence:

$$c_1^B(e_2, \tau) > c_{2\beta}^B(\tau;\, [e_2, e_H])$$

By repeating the above procedure for e_2 we obtain that there must exist e_3 s.t. $e_L < e_3 < e_2$ and

$$c_1^B(e_3, \tau) > c_{2\beta}^B(\tau;\, [e_3, e_H])$$

Therefore we can construct a decreasing sequence, $\{e_i\}_{i=1}^{\infty}$, bounded from below by e_L and hence convergent to e_L:

$$\underset{i \to \infty}{lim}\, e_i = e_L$$

where for each i we have:

$$\xi\left(e_i\right) \equiv c_1^B(e_i, \tau) - c_{2\beta}^B(\tau;\, [e_i, e_H]) > 0$$

Taking the limit of $\{\xi\left(e_i\right)\}_{i=1}^{\infty}$ to obtain:

$$\underset{i\to\infty}{lim}\xi\left(e_i\right) \to \xi\left(e_L\right) = c_1^B(e_L, \tau) - c_{2\beta}^B(\tau;\, [e_L, e_H]) < 0$$

Therefore there must exist $\hat{i}$ such that for $i > \hat{i}$ we have $\xi\left(e_i\right) < 0$ - a contradiction since we must also have $\xi\left(e_i\right) > 0$ for each i. As a result, we conclude that our initial assumption $c_1^B(e_1, \tau) \geq c_{2\beta}^B(\tau;\, [e_1, e_H])$ cannot be true and we must have:

$$c_1^B(e_1, \tau) < c_{2\beta}^B(\tau;\, [e_1, e_H])$$

That is, for each e_1 with $e_1 < e^F(\tau)$, the panic set $[e_1, e_H]$ cannot be part of equilibrium.

Step 2 Consider $P(\tau) \subseteq [e_L, e_H]$ and define:

$$e^* \equiv \{e \in P(\tau)\; s.t.\; e \leq z\; for\; all\; z \in P(\tau)\}$$

If $e^* < e^T(\tau)$, by step 1 the panic set $[e^*, e_H]$ cannot be part of equilibrium and we must have $P(\tau) \subset [e^*, e_H]$. But, applying lemma 2, we also obtain that:

$$c_1^B(e^*, \tau) < c_{2\beta}^B(\tau;\, [e^*, e_H]) \leq c_{2\beta}^B(\tau;\, P(\tau))$$

hence $P(\tau)$ cannot be part of equilibrium whenever there is e such that $e \in P(\tau)$ and $e < e^T(\tau)$. □

Proposition 1.9.

Proof. See the discussion in the text. □

Proposition 1.10.

Proof. (i) We want to show that for each $\tau \in (0,1)$ there exist a wealth gap $\Delta^* \in [0,\ f^{-1})$ such that for $\Delta \in (\Delta^*, f^{-1})$ the equilibrium where the low-income investors follow the panic strategy does not exist. That is, we must show that when $\Delta > \Delta^*$ those with low-income will not be best responding with the panic strategy in (1.2):

$$c_1(e_L(\Delta), \tau) < c_{2\beta}(e_L(\Delta), \tau)$$

For given values of τ and Δ, let $b^*(e_L(\Delta), \tau)$ be defined as the solution to:

$$c_1^B(e_L(\Delta), \tau) = \frac{R^{\frac{1}{\gamma}}\left((1-\tau)e_L(\Delta) - \pi c_1^B(e_L(\Delta), \tau) + b^*(e_L(\Delta), \tau)\right)}{(1-\pi)\left(\pi + (1-\pi)R^{\frac{1-\gamma}{\gamma}})\right)}$$

From (1.10) - (1.11) and (1.22), the function $f_L(\Delta)$:

$$f_L(\Delta) \equiv (1-\tau)e_L(\Delta) - \pi c_1(e_L(\Delta), \tau)$$

is decreasing in Δ. In addition, $c_1^B(e_L(\Delta), \tau)$ is also decreasing in Δ and hence we have:

$$c_1^B(e_L(\Delta), \tau) < c_{2\beta}^B(e_L(\Delta), \tau) \;\Leftrightarrow\; b(e_L(\Delta), \tau) > b^*(e_L(\Delta), \tau)$$

From, (1.13) - (1.16), and the fact that $f_L(\Delta)$ is decreasing in Δ, we also obtain:

$$\frac{\partial b(e_L(\Delta), \tau)}{\partial \Delta} > 0$$

The above implies that there exist $\tilde{\Delta} \in [0,\ f^{-1})$ such that for $\Delta > \tilde{\Delta}$ banks servicing those with low-income best respond by qualifying for a bailout, i.e. $W(e_L(\Delta), \tau) > 0$ for $\Delta > \tilde{\Delta}$. Finally, for $\Delta \to f^{-1}$ we have $f_L(\Delta) \to 0$ and therefore $b^*(e_L(\Delta), \tau) \to 0$. Hence, there must exist a level of inequality $\Delta^{**} \in [0,\ f^{-1})$ such that for $\Delta > \Delta^* \equiv max\left\{\Delta^{**}, \tilde{\Delta}\right\}$ we have $b^*(e_L(\Delta), \tau) < b(e_L(\Delta), \tau)$, which in turn implies that $c_1^B(e_L(\Delta), \tau) < c_{2\beta}^B(e_L(\Delta), \tau)$. That is, when the wealth gap is sufficiently large, banks servicing investors with low-income would best respond by qualifying for a bailout and, at the same time, the anticipated bailout would prevent a panic for the low-income investors from taking place in equilibrium. As a result, an equilibrium run on the banks cannot involve those with low-income when $\Delta > \Delta^*$.

(ii) I will show that for each value of the tax rate τ in the interval $(0, 1)$ we can find $\hat{f} > 0$ and $\hat{\Delta} \in [0, \hat{f}^{-1})$ such that for $f < \hat{f}$ and for $\Delta > \hat{\Delta}$ there exists an equilibrium where the high-income investors follow the panic strategy. First, by an argument analogous to part (i), we can show that for each τ, the bailout to banks providing intermediation to investors with high-income is strictly decreasing as a function of the wealth gap Δ:

$$\frac{\partial b\left(e_H\left(\Delta\right), \tau\right)}{\partial \Delta} < 0$$

Second, let $e^{NB}\left(\tau\right)$ be such that $W\left(e^{NB}\left(\tau\right), \tau\right) < 0$ for $e > e^{NB}\left(\tau\right)$. That is, when the tax rate is equal to τ, those banks servicing investors whose endowment exceeds $e^{NB}\left(\tau\right)$ will choose to self-insure in case their investors follow the panic strategy. For each $e^{NB}\left(\tau\right)$ there must exist $\hat{f} > 0$ such that for all $f < \hat{f}$ we have:

$$e^{NB}\left(\tau\right) < 1 + \frac{(1-f)}{f}$$

therefore for a wealth gap Δ in the range $\left(\hat{\Delta}, \dfrac{1}{f}\right)$, where $\hat{\Delta}$ is given by:

$$\hat{\Delta} = \frac{e^{NB}\left(\tau\right) - 1}{1-f}$$

we would have:

$$e_H(\Delta) = 1 + (1-f)\Delta > e^{NB}\left(\tau\right)$$

and therefore

$$W\left(e_H(\Delta),\tau\right) < 0$$

that is, for all $f < \hat{f}$ and for all $\Delta > \hat{\Delta}$ banks servicing those with high-income choose to self-insure. Next, from (1.25) we obtain that high-income investors are best responding with the panic strategy when their banks are not bailed out. Therefore, when the level of inequality is sufficiently high, the equilibrium where the high-income investors run on the banks would exists. □

Proposition 1.11.

Proof. Follows from the discussion in the text. □

Proposition 1.13.

Proof. Suppose that Ω^* is a stable coalition structure and let $\{C_k^*(\tau)\}_{k=0}^1$ be the collection of payment schedules generated by Ω^*. I show each $\Omega_k^* \in \Omega^*$ is not blocked only if for all $e \in \Omega_k^*$ we have $C_k^*(e,\tau) = C_S(e,\tau)$. That is, a stable coalition structure is payoff-equivalent to type specific coalition structure . There are two cases to consider. In case 1, the coalition Ω_k^* is not bailed out, whereas in case 2, coalition Ω_k^* is bailed out.

Case 1: No-bailout I derive the payment schedule for each $e \in \Omega_k^*$ under general set of weights and then show how to pick these weight in order to ensure that the proportional rule will be satisfied. Denote with $\omega_k(e)$ the weight on type e investors in coalition k and with $g_k(e)$ - the fraction of type e investors in the coalition. In state s the coalition distributes its remaining resources to maximize:

$$V_s^k\left(\psi_s^k(\tau)\right) \equiv \int_{e_L}^{e_H} g_k(e)\,\omega_k(e)\left\{(1-\pi)\left[\pi_s(e)u\left(c_{1s}^k(e,\tau)\right) + (1-\pi_s(e))\,u\left(c_{2s}^k(e,\tau)\right)\right]\right\} de \tag{1.50}$$

subject to the budget constraint in state s:

$$\int_{e_L}^{e_H} g_k\,(e) \left\{(1-\pi)\left[\pi_s(e)c_{1s}^k(e,\tau) + (1-\pi_s(e))\,\frac{c_{2s}^k(e,\tau)}{R}\right]\right\} de = \psi_s^k(\tau) \tag{1.51}$$

and where $\psi_s^k(\tau)$ - the the remaining resources in state s - are equal to:

$$\psi_s^k(\tau) = \int_{e_L}^{e_H} g_k\,(e)\left[(1-\tau)e - \pi c_1^k(e,\tau)\right] de \tag{1.52}$$

The first order condition for type e investors becomes:

$$u'(c_{1s}^k(e,\tau)) = Ru'(c_{2s}^k(e,\tau)) = \frac{\mu_s^k(\tau)}{\omega_k(e)} \tag{1.53}$$

where $\mu_s^k(\tau)$ denotes the shadow on the budget constraint in state s. On the other hand, the payments during the first π withdrawals will be set to maximize:

$$\int_{e_L}^{e_H} g_k\,(e)\,\omega_k(e)\left\{\pi u\left(c_1^k(e,\tau)\right)\right\} de + (1-q)V_\alpha^k\left(\psi_\alpha^k(\tau)\right) + qV_\beta^k\left(\psi_\beta^k(\tau)\right) \tag{1.54}$$

and we obtain the following first order condition for type e investors:

$$u'(c_1^k(e,\tau)) = (1-q)\frac{\mu_\alpha^k(\tau)}{\omega_k(e)} + q\frac{\mu_\beta^k(\tau)}{\omega_k(e)} \tag{1.55}$$

Next, I use the functional form for u (namely, constant relative risk aversion with $\gamma > 1$) in order to derive an explicit solution for the payment schedule. In particular, the weights are normalized to satisfy:

$$\int_{e_L}^{e_H} g_k\,(e)\,(w_k(e))^{\frac{1}{\gamma}}\,de = 1 \tag{1.56}$$

and we obtain that the payment schedule for each investor type in the coalition is equal to:

$$c_1^k(e,\tau) = (w_k(e))^{\frac{1}{\gamma}} (1-\tau)\, \hat{e}_k \frac{\lambda_k^{NB}}{1+\pi\lambda_k^{NB}} \tag{1.57}$$

$$c_{1s}^k(e,\tau) = (w(e))^{\frac{1}{\gamma}} \left(\int_{e_L}^{e_H} g_k\,(e)\,\lambda_s(e)de\right) \left(\frac{\psi_s^k(\tau)}{1-\pi}\right) \tag{1.58}$$

$$c_{2s}^k(e,\tau) = R^{\frac{1}{\gamma}} c_{1s}^k(e,\tau) \tag{1.59}$$

where

$$\hat{e}_k = \int_{e_L}^{e_H} g_k\,(e)\, e de \tag{1.60}$$

$$\lambda_k^{NB} = \frac{1}{1-\pi}\left\{(1-q)\left(\int_{e_L}^{e_H} g_k\,(e)\,\lambda_\alpha(e)de\right)^{-\gamma} + q\left(\int_{e_L}^{e_H} g_k\,(e)\,\lambda_\beta(e)de\right)^{-\gamma}\right\}^{-\frac{1}{\gamma}} \tag{1.61}$$

$$\lambda_s(e) = \frac{1}{\pi_s(e) + (1-\pi_s(e))R^{1-\frac{1}{\gamma}}} \tag{1.62}$$

$$\pi_\alpha(e) = 0 \quad for \quad e \in [e_L, e_H] \tag{1.63}$$

$$\pi_\beta(e) = \left\{ \begin{array}{c} 0 \\ \pi \end{array} \right\} if \left\{ \begin{array}{c} e \notin P(\tau) \\ e \in P(\tau) \end{array} \right\} \tag{1.64}$$

For $\{e_1, e_2\} \subseteq \Omega_k^*$, the proportional rule requires:

$$\frac{c_1^k(e_1,\tau)}{c_1^k(e_2,\tau)} = \frac{c_{2\alpha}^k(e_1,\tau)}{c_{2\alpha}^k(e_2,\tau)} = \frac{c_{1\beta}^k(e_1,\tau)}{c_{1\beta}^k(e_2,\tau)} = \frac{c_{2\beta}^k(e_1,\tau)}{c_{2\beta}^k(e_2,\tau)} = \frac{e_1}{e_2}$$

In order to satisfy this restriction, the coalition must assign weights to different type of investors according to:

$$w_k(e) = \left(\frac{e}{\hat{e}_k}\right)^{\gamma} \tag{1.65}$$

From (1.56) - (1.59) and (1.65) we obtain that the payment schedule for type e investors in the coalition becomes:

$$c_1^k(e,\tau) = (1-\tau)\, e \frac{\lambda_k^{NB}}{1+\pi\lambda_k^{NB}} \tag{1.66}$$

$$c_{1s}^k(e,\tau) = \left(\int_{e_L}^{e_H} g_k\left(e\right)\lambda_s(e)de\right)\left(\frac{(1-\tau)\, e}{(1-\pi)\left(1+\pi\lambda_k^{NB}\right)}\right) \tag{1.67}$$

$$c_{2s}^k(e,\tau) = R^{\frac{1}{\gamma}}\left(\int_{e_L}^{e_H} g_k\left(e\right)\lambda_s(e)de\right)\left(\frac{(1-\tau)\, e}{(1-\pi)\left(1+\pi\lambda_k^{NB}\right)}\right) \tag{1.68}$$

First, in the special case where all investors in the coalition have the same endowment $\{e\} = \Omega_k^*$, we have $w_k(e) = 1$. The conditions characterizing the payment schedule in this case will be the same as (1.10), (1.11) and (1.23) and we have $C_k^*\left(e,\tau\right) = C_S\left(e,\tau\right)$. Second, if all investors in the coalition follow the same withdrawal strategy, (1.60) - (1.68) would imply that for each $e \in \Omega_k^*$ we have $C_k^*\left(e,\tau\right) = C_S\left(e,\tau\right)$. Third, if a fraction of the investors in Ω_k^* follow the panic strategy, then (1.60) - (1.68) imply that each investor type that follows the no-panic withdrawal strategy and belongs to Ω_k^* will be strictly better off by forming their own coalition since for $e \notin P(\tau)$ we have $C_k^*\left(e,\tau\right) \ll C_S\left(e,\tau\right)$ and therefore $W_k(e,\tau) < W_S(e,\tau)$. Thus, any coalition with $b_k(\tau) = 0$ mixing investors that follow the panic withdrawal strategy with investors that follow the no-panic withdrawal strategy will be blocked and as a result cannot belong to a stable coalition structure. We conclude that for $e \in \Omega_k^*$ such that $b_k(\tau) = 0$ we have $C_k^*\left(e,\tau\right) = C_S\left(e,\tau\right)$.

Case 2: Bailout When $b_k(\tau) > 0$, then for $e \in \Omega_k^*$ the payment schedule for type e investors in state β - $\left(c_{1\beta}^k(e,\tau), c_{2\beta}^k(e,\tau)\right)$ - will be set by the government according to (1.40) - (1.42). We want to show that for each $e \in \Omega_k^*$, we have $C_k\left(e,\tau\right) = C_S\left(e,\tau\right)$, that is, the payment schedule for the investors in Ω_k^* is payoff-equivalent to a separate coalition for each investor type in Ω_k^*. So, define $\hat{c}_{2\alpha}(e,\tau)$ as:

$$\hat{c}_{2\alpha}(e,\tau) \equiv \frac{R\left((1-\tau)e - \pi c_1^k\left(e,\tau\right)\right)}{(1-\pi)} \tag{1.69}$$

Note that if there exist $e \in \Omega_k^*$ such that

$$c_{2\alpha}^k\left(e,\tau\right) > \hat{c}_{2\alpha}(e,\tau) \tag{1.70}$$

there must also exist $e' \in \Omega_k^*$ such that:

$$c_{2\alpha}^k\left(e',\tau\right) < \hat{c}_{2\alpha}(e',\tau) \tag{1.71}$$

otherwise, we have:

$$\int_{e_L}^{e_H} g_k\left(e\right)\left[c_{2\alpha}^k(e,\tau) - \frac{R\left[(1-\tau)e - \pi c_1^k(e,\tau)\right]}{(1-\pi)}\right] de > 0 \tag{1.72}$$

which would violate the budget constraint in state α. At the same time, if there is $e' \in \Omega_k^*$ such that (1.71) holds then Ω_k^* will be blocked. The reason is that type e' would strictly prefer to form their own coalition. To see this, note from (1.69) that $\left(c_1^k\left(e',\tau\right), \hat{c}_{2\alpha}(e',\tau)\right)$ is feasible when type e' operate their own coalition. In addition, $c_{2\alpha}^k\left(e',\tau\right) < \hat{c}_{2\alpha}(e',\tau)$ combined with (1.40) and (1.41) yields:

$$\left(c_{1\beta}^S(e',\tau), c_{2\beta}^S(e',\tau)\right) \geq \left(c_{1\beta}^k(e',\tau), c_{2\beta}^k(e',\tau)\right)$$

Thus we have

$$\left(c_1^k\left(e',\tau\right), c_{1\beta}^S(e',\tau), c_{2\beta}^S(e',\tau)\right) \geq \left(c_1^k\left(e',\tau\right) c_{1\beta}^k(e',\tau), c_{2\beta}^k(e',\tau)\right)$$

and $c_{2\alpha}^k\left(e',\tau\right) < \hat{c}_{2\alpha}(e',\tau)$, which would imply that type e' are better off on their own. Therefore, if Ω_k^* is not blocked, we must have for each $e \in \Omega_k^*$:

$$c_{2\alpha}^{k}(e,\tau) = \frac{R\left((1-\tau)e - \pi c_1^k\,(e,\tau)\right)}{(1-\pi)} \tag{1.73}$$

Note that for each $e \in \Omega_k^*$ a separate bank for type e will be able to exactly replicate $C_k\,(e,\tau)$ by setting $\left(c_1^k(e,\tau), c_{2\alpha}^k(e,\tau)\right) = \left(c_1^S(e,\tau), c_{2\alpha}^S(e,\tau)\right)$. Indeed, $\left(c_1^k(e,\tau), c_{2\alpha}^k(e,\tau)\right)$ must satisfy (1.72) and therefore is feasible when type e operate on their own. In addition, from (1.40) and (1.41) we obtain that if

$$\left(c_1^S(e,\tau), c_{2\alpha}^S(e,\tau)\right) = \left(c_1^k(e,\tau), c_{2\alpha}^k(e,\tau)\right)$$

we must also have:

$$\left(c_{1\beta}^S(e,\tau), c_{2\beta}^S(e,\tau)\right) = \left(c_{1\beta}^k(e,\tau), c_{2\beta}^k(e,\tau)\right)$$

But the optimal payment schedule for type e investors among all those satisfying (1.72) will coincide with the optimal payment schedule when this type operate on their own, namely $C_S\,(e,\tau)$. Therefore if Ω_k^* is not blocked, we must have $C_k\,(e,\tau) = C_S\,(e,\tau)$ for each $e \in \Omega_k^*$ as desired. $\square$

Chapter 2

Bailouts, Bail-ins and Banking Crises

2.1 Introduction

In the years since the financial crisis of 2008 and the associated bailouts of banks and other financial institutions, policy makers in several jurisdictions have drafted rules requiring that these institutions impose losses on (or "bail in") their investors in any future crisis. These rules aim both to protect taxpayers in the event of a future crisis and to change the incentives of banks and investors in a way that makes a crisis less likely. While the specific requirements vary, and are often yet to be finalized, in many cases the bail-in will be triggered by an announcement or action taken by the bank itself. This fact raises the question of what incentives banks will face when deciding whether and when to take actions that bail in their investors. In this paper, we ask how the prospect of being bailed out by the government influences banks' bail-in decisions and how these decisions, in turn, affect the susceptibility of the banking system to a run by investors.

At one level, the reason why banks and other financial intermediaries sometimes experience runs by their investors is well understood. Banks offer deposit contracts that allow investors to withdraw their funds at face value on demand or at very short notice. During a bank run, investors fear that a combination of real losses and/or heavy withdrawals will leave their bank unable to meet all of its obligations. This belief makes it individually rational for each investor to withdraw her funds at the first opportunity; the ensuing rush to withdraw then guarantees that the bank does indeed fail, justifying investors' pessimistic beliefs.[1]

[1]This basic logic applies not only to commercial banking to also to a wide range of financial

A key element of this well-known story is that the response to a bank's losses and/or a run by its investors is *delayed*. In other words, there is a period of time during which a problem clearly exists and investors are rushing to withdraw, but the bank continues to operate as normal. Only when the situation becomes bad enough is some action – freezing deposits, renegotiating obligations, imposing losses on some investors, etc. – taken. This delay tends to deepen the crisis and thereby increase the incentive for investors to withdraw their funds at the earliest opportunity.

From a theoretical perspective, this delayed response to a crisis presents something of a puzzle. A run on the bank creates a misallocation of resources that makes the bank's investors as a group worse off. Why do these investors not collectively agree to an alternative arrangement that efficiently allocates whatever losses have occurred while minimizing liquidation and other costs? In particular, why does the banking arrangement not respond more quickly to whatever news leads investors to begin to panic and withdraw their funds?

Most of the literature on bank runs resolves this puzzle using an incomplete-contracts approach.[2] In particular, it is typically assumed to be impossible to write and/or enforce the type of contracts that would be needed to generate fully state-contingent payments to investors. The classic paper of Diamond and Dybvig (1983), for example, assumes that banks must pay withdrawing investors at face value until the bank has liquidated all of its assets and is completely out of funds. Other contracts – in which, for example, the bank is allowed to impose withdrawal fees when facing a run – are simply not allowed. Even those more recent papers that study more flexible banking arrangements impose some incompleteness of contracts. Peck and Shell (2003), for example, allow a bank to adjust payments to withdrawing investors

intermediation arrangements. See Yorulmazer (2014) for a discussion of a several distinct financial intermediation arrangements that experienced run-like episodes during the financial crisis of 2008.

[2]An important exception is Calomiris and Kahn (1991), in which the ex post misallocation of resources associated with a run is part of a desirable ex ante incentive arrangement to disciple bankers' behavior.

based on any information it receives. However, the bank is assumed not to observe the realization of a sunspot variable that is available to investors and, in this sense, the ability to make state-contingent payments is still incomplete.[3]

If the fundamental problem underlying the fragility of banking arrangements is incompleteness of contracts, then an important goal of financial stability policy should be to remove this incompleteness. In other words, a key conclusion of the literature to date is that policy makers should aim to create legal structures under which more fully state-contingent banking contracts become feasible. There has, in fact, been substantial progress in this direction in recent years, including the establishment of orderly resolution mechanisms for large financial institutions and other ways of "bailing in" these institutions' investors more quickly and more fully than in the past. The reform of money market mutual funds that was adopted in the U.S. in 2014 is a prime example. Under the new rules, these funds are permitted to temporarily prohibit redemptions (called "erecting a gate") and impose withdrawal fees during periods of high withdrawal demand if doing so is deemed to be in the best interests of the funds' investors.

In this paper, we ask whether making banking arrangements more fully state contingent – thereby allowing banks increased flexibility to bail in their investors – is sufficient to eliminate the problem of bank runs. To answer this question, we study a model in the tradition of Diamond and Dybvig (1983), but in which banks can freely adjust payments to investors based on any information available to the bank or to its investors. We think of this assumption as capturing an idealized situation in which policy makers' efforts to improve the contractual environment have been completely successful. We ask whether and under what conditions bank runs can occur in this idealized environment.

[3] This same approach is taken in a large number of papers that study sunspot-driven bank runs in environments with flexible banking contracts, including Ennis and Keister (2010), Sultanum (2015), Keister (2016), and many others. See Andolfatto et al. (2016) for an interesting model in which the bank does not observe the sunspot state, but can attempt to elicit this information from investors.

There are two aggregate states in our model and banks face uncertainty about the value of their investments. No banks experience losses in the good aggregate state, but in the bad aggregate state, some banks' assets are impaired. The government is benevolent and taxes agents' endowments in order to provide a public good. If there is a banking crisis, the government can also use these resources to provide bailouts to impaired banks. The government observes the aggregate state but cannot immediately tell which banks have impaired assets and which do not. In addition, the government cannot commit to a bailout plan; instead, the payment made to each bank will be chosen as a best response to the situation at hand. As in Keister (2016), this inability to commit implies that banks in worse financial conditions will receive larger bailout payments, as the government will aim to equalize the marginal utility of consumption across agents to the extent possible.

A bank with impaired assets has fewer resources available to make payments to investors. In an efficient allocation, such a bank would respond by immediately bailing in its investors, reducing all payments so that the loss is evenly shared. When the bank anticipates a government intervention, however, it may have an incentive to delay this response. By instead acting as if its assets were not impaired, the first group of its depositors who withdraw will receive higher payments. The government will eventually learn that the bank's assets are impaired and, at this point, will find the bank to be in worse financial shape as a result of the delayed response. The inability to commit prevents the government from being able to punish the bank at this point; instead, the bank will be given a larger bailout payment as the government aims to raise the consumption levels of its remaining investors. This larger payment then justifies the bank's original decision to delay taking action. In other words, we show that *bailouts delay bail-ins*.

The delay in banks' bail-in decisions has implications at both the aggregate and bank level. The delayed response makes banks with weak fundamentals even worse off and leads the government to make larger bailout payments, at the cost of a lower level of public good provision for everyone. In some cases, the misallocation of resources created by the delay may be large enough to give investors in weak banks an incentive to *run* in an attempt to withdraw before the bail-in is enacted. In these cases, the delayed bail-in creates financial fragility.

Our approach has novel implications for the form a banking crisis must take. Models in the tradition of Diamond and Dybvig (1983) typically do not distinguish between a single bank and the banking system; one can often think of the same model as applying equally well to both situations. If the banking system is composed of many banks, such models predict that there could be a run on a single bank, on a group of banks, or on all banks, depending on how each bank's depositors form their beliefs. In our model, in contrast, there cannot be a run on only one bank, nor can there be a crisis in which only one bank chooses to delay bailing in its investors. If there is only a problem at one bank in our model, the government will choose to provide full deposit insurance, which removes any incentive for investors to run as well as any need for the bank to enact a bail-in. The problems of bank runs and delayed bail-ins can only arise in this model if the underlying losses are sufficiently widespread. In addition, a bank run in this model cannot be driven purely by sunspots. If a bank's assets are not impaired, we show that there is no incentive for it to delay its response should a run occur. Given that there is no delay in the response, the bank's depositors will have no incentive to run.

The remainder of the paper is organized as follows. The next section describes the economic environment and the actions available to banks, investors, and the government. In Section 2.3, we derive the constrained efficient allocation of resources in this environment, which is a useful benchmark for what follows. In Section 2.4, we analyze equilibrium in the withdrawal game played by an individual bank's investors, taking actions elsewhere in the economy as given. We study equilibrium in the overall economy in Section 2.5 before concluding in Section 2.6.

2.2 The model

We base our analysis on a version of the Diamond and Dybvig (1983) model with flexible banking contracts and fiscal policy conducted by a government with limited commitment. We introduce idiosyncratic risk to banks' asset holdings and highlight how banks' incentives to react to a loss are influenced by their anticipation of government intervention. In this section, we introduce the agents, preferences, and technologies that characterize the economic environment.

2.2.1 The environment

Time. There are three time periods, labeled $t = 0, 1, 2$.

Investors. There is a continuum of investors, indexed by $i \in [0,1]$, in each of a continuum of locations, indexed by $k \in [0,1]$. Each investor has preferences characterized by

$$U\left(c_1^{i,k}, c_2^{i,k}, g; \omega^{i,k}\right) \equiv u(c_1^{i,k} + \omega^{i,k} c_2^{i,k}) + v(g), \tag{2.1}$$

where $c_t^{i,k}$ denotes the period-t private consumption of investor i in location k and g is the level of the public good, which is available in all locations. The random variable $\omega^{i,k} \in \Omega \equiv \{0,1\}$ is realized at $t = 1$ and is privately observed by the investor. If $\omega^{i,k} = 0$, she is *impatient* and values private consumption only in period 1, whereas if $\omega^{i,k} = 1$ she values consumption equally in both periods. Each investor will be impatient with a known probability $\pi > 0$, and the fraction of investors who are impatient in each location will also equal π. The functions u and v are assumed to be smooth, strictly increasing, strictly concave and to satisfy the usual Inada conditions. As in Diamond and Dybvig (1983), the function u is assumed to exhibit a coefficient of relative risk aversion that is everywhere greater than one. Each investor is endowed with one unit of of an all-purpose good at the beginning of period 1 and nothing in subsequent periods. Investors cannot directly invest their endowments and must instead deposit with a financial intermediary.

Banks. In each location, there is a representative financial intermediary that we refer to as a *bank*.[4] Each bank accepts deposits in period 0 from investors in its location and invests these funds in a set of ex ante identical projects. A project requires one unit of input at $t = 0$ and offers a gross return of 1 at $t = 1$ or of $R > 1$ at $t = 2$ if it is not impaired. In period 1, however, $\sigma_k \in \Sigma \equiv \{0, \bar{\sigma}\}$ of the projects held by bank k will be revealed to be impaired. An impaired project is worthless: it produces zero return in either period. We will refer to σ_k as the *fundamental state* of bank k. A bank with $\sigma_k = 0$ is said to have *sound* fundamentals, whereas a bank with $\sigma_k = \bar{\sigma}$ is said to have *weak* fundamentals. The realization of σ_k is observed at the beginning of $t = 1$ by the bank's investors, but is not observed by anyone outside of location k.

After investors' preference types and banks' fundamental states are realized, each investor informs her bank whether she wants to withdraw in period 1 or in period 2. The bank observes *all* reports from its investors *before* making any payments to withdrawing investors. Those investors who chose to withdraw in period 1 then begin arriving sequentially at the bank in a randomly-determined order. Investors are isolated from either other during this process and no trade can occur among them; each investor simply consumes the payment she receives from her bank and returns to isolation. As in Wallace (1988) and others, this assumption prevents re-trading opportunities from undermining banks' ability to provide liquidity insurance.

Aggregate uncertainty. The fraction of banks whose assets are impaired depends on the aggregate state of the economy, which is either *good* or *bad*. In the good state, all banks have sound fundamentals. In the bad state, in contrast, a fraction $n \in [0, 1]$ of banks have weak fundamentals and, hence, total losses in the financial system

[4]While we use the term "bank" for simplicity, our model should be interpreted as applying to the broad range of financial institutions that engage in maturity transformation.

are $n\bar{\sigma}$. The probability of the bad state is denoted q; we interpret this event as an economic downturn that has differing effects across banks. If we think of the projects in the model as representing loans, for example, then the loans made by some banks are relatively unaffected by the downturn (for simplicity, we assume they are not affected at all), while other banks find they have substantial non-performing loans. Conditional on the bad aggregate state, all banks are equally likely to experience weak fundamentals. The ex-ante probability that a given bank's fundamentals will be weak is, therefore, equal to qn.

The government. The government in our model acts as both a fiscal authority and a banking supervisor. Its objective is to maximize the sum of all investors' expected utilities at all times. The government's only opportunity to raise revenue comes in period 0, when it chooses to tax investors' endowments at rate τ. In period 1, the government will use this revenue to provide the public good and, perhaps, to make transfers (*bailouts*) to banks. The government is unable to commit to the details of the bailout intervention ex-ante, but instead chooses the policy ex post, as a best response to the situation at hand.

The government observes the aggregate state of the economy at the beginning of period but, when the aggregate state is bad, is initially unable to determine *which* banks have weak fundamentals. After a measure $\theta \geq 0$ of investors have withdrawn from each bank, the government observes the idiosyncratic state σ_k of all banks and decides how to allocate its tax revenue between bailout payments to banks and the public good. The parameter θ thus measures how quickly the government can collect bank-specific information during a crisis and respond to this information. Banks that receive a bailout from the government are immediately placed in *resolution* and all subsequent payments made by these banks are chosen by the government. Once the public good has been provided, the government no longer has access to any resources and there will be no further bailouts.

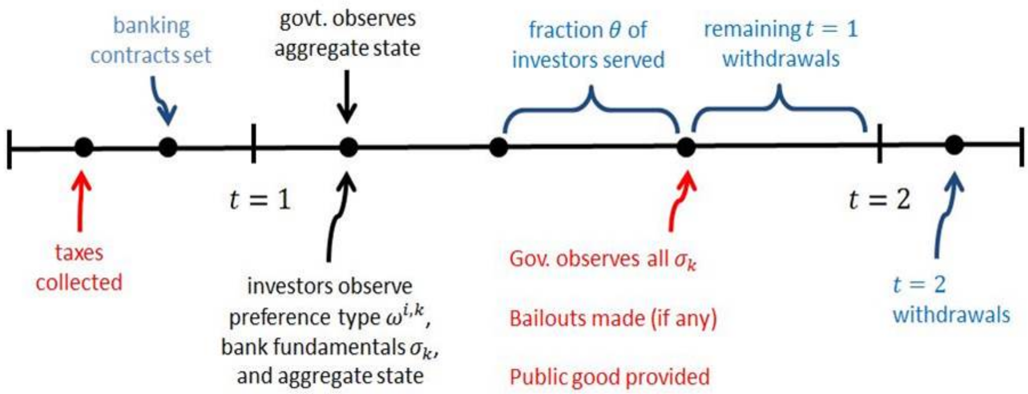


Figure 2.1: Sequence of events

2.2.2 Timeline

The sequence of events is depicted in Figure 2.1. In period 0, the government chooses the tax rate τ on endowments and investors deposit their after-tax endowment with the bank in their location. At the beginning of period 1, each investor observes her own preference type and the fundamental state of her bank; she then decides whether to withdraw in period 1 or in period 2. Banks observe the choices of their investors and begin making payments to withdrawing investors as they arrive. Once the measure of withdrawals reaches θ, the government observes all banks' fundamental states. At this point, the government may choose to bail out banks with weak fundamentals and places any banks that were bailed out into resolution. After bailout payments are made, all remaining tax revenue is used to provide the public good. Banks that were not bailed out out continue to make payments to investors according to their contract, while the remaining payments made by banks in resolution are dictated by the government.

Discussion

While our model contains many elements that are familiar from the literature on bank runs, there are some key differences. Perhaps most importantly, banks in our model are able to condition payments to *all* investors on the *total* demand for early withdrawal. Green and Lin (2003) refer to this assumption as "the case without sequential service." This language is potentially confusing when applied to our model: banks still serve withdrawing investors sequentially here. The key point, however, is that a bank is able to observe early withdrawal demand *before* deciding how to allocate resources across agents. By allowing all payments made by the bank to depend on this information, our contract space is larger than that in most of the bank runs literature. In taking this approach, we aim to capture a contractual environment that is sufficiently rich to eliminate the underlying sources of bank runs that appear in the existing literature.

The role of aggregate uncertainty in our model is to force the government to fix a tax plan before knowing the aggregate losses of the banking system. If the government knew in advance how many banks would experience loses, it would collect additional taxes at $t = 0$ for the purpose of providing insurance against this location-specific shock. In fact, given that we assume the government can costlessly raise revenue through lump-sum taxes, it would collect enough revenue to provide complete insurance. Our timing assumption makes providing this insurance costly. If, for example, the probability q of the bad state is close to zero, the government will collect tax revenue equal to the desired level of the public good in the good aggregate state. If the realized state turns out to be bad, the marginal value of public resources will increase, but the government will be unable to raise additional revenue.

2.3 The constrained efficient allocation

We begin by studying an allocation that will serve as a useful benchmark in the analysis. Suppose a benevolent planner could control the operations of all banks and the government, as well as investors' withdrawal decisions. This planner observes all of the information available to banks and investors, including each investor's preference type. It faces the same restrictions on fiscal policy as the government; in particular, all tax revenue must be raised at $t = 0$, before the aggregate state is realized. The planner allocates resources to maximize the sum of all investors' utilities.

It is fairly easy to see that the planner will direct all impatient investors to withdraw at $t = 1$, since they do not value later consumption, and will direct all patient investors to withdraw at $t = 2$, since it is less expensive to provide consumption to them after investment has matured. In addition, because investors are risk averse, the planner will choose to treat investors and banks symmetrically. In the good aggregate state, the planner will give a common level of consumption c_{10} in period 1 to all impatient investors and a common level c_{20} in period 2 to all patient investors. (The second subscript indicates that these consumption levels pertain to the good aggregate state, where zero banks have weak fundamentals.) In the bad aggregate state, the planner will give a common consumption profile (c_{1S}, c_{2S}) to investors in all banks with strong fundamentals and a common profile $(c_{1W}, _{2W})$ to investors in all banks with weak fundamentals. These consumption levels will be chosen to maximize

$$
\begin{aligned}
&(1-q)\left\{\pi u\left(c_{10}\right)+(1-\pi)\, u\left(c_{20}\right)+v\left(\tau\right)\right\}\\
&+q\left\{\begin{array}{c}
(1-n)\left(\pi u\left(c_{1S}\right)+(1-\pi)\, u\left(c_{2S}\right)\right)+n\left(\pi u\left(c_{1W}\right)+(1-\pi)\, u\left(c_{2W}\right)\right)\\
+v\left(\tau-(1-n)b_S-nb_W\right)
\end{array}\right\}.
\end{aligned}
$$

subject to feasibility constraints

$$\pi c_{10} + (1-\pi)\frac{c_{20}}{R} \leq 1-\tau$$
$$\pi c_{1S} + (1-\pi)\frac{c_{2S}}{R} \leq 1-\tau+b_S$$
$$\pi c_{1W} + (1-\pi)\frac{c_{2W}}{R} \leq 1-\tau-\bar{\sigma}+b_W,$$

where b_z denotes the per-investor transfer (or "bailout") given to each bank of type z in the bad aggregate state.[5] The three constraints each state that the present value of the consumption given to depositors in a bank must come from the initial deposit $1-\tau$, minus the loss $\bar{\sigma}$ for banks with weak fundamentals, plus any bailout received. The restriction that the planner cannot raise additional tax revenue in period 1 is equivalent to saying that the bailout payments must be non-negative,

$$b_S \geq 0 \quad \text{and} \quad b_W \geq 0. \tag{2.2}$$

The first-order conditions for the optimal consumption levels can be written as

$$u'(c_{1z}) = Ru'(c_{2z}) = \mu_z \qquad \text{for } z = 0, S, W, \tag{2.3}$$

where μ_z is the Lagrange multiplier on the resource constraint associated with state z normalized by the probability of a bank ending up in that state. The first-order condition for the choice of tax rate τ can be written as

$$(1-q)v'(\tau) + qv'(\tau-(1-n)b_S - nb_W) = (1-q)\mu_0 + q(1-n)\mu_S + qn\mu_W, \tag{2.4}$$

which states that the expected marginal value of a unit of public consumption equals

[5]Note that our notation does not allow the planner to make bailout payments in the good aggregate state. This assumption prevents the planner from being able to make tax revenue fully state-contingent by, for example, setting $\tau = 1$ and holding all resources outside of the banking system until the aggregate state is revealed.

the expected marginal value of a unit of private consumption at $t = 0$. The first-order conditions for the bailout payments are

$$v'(\tau - (1-n)b_S - nb_W) \geq \mu_z, \quad \text{with equality if } b_z > 0, \quad \text{for } z = S, W \tag{2.5}$$

If the marginal value of private consumption in some banks were higher than the marginal value of public consumption in the bad aggregate state, the planner would transfer resources to (or "bail out") these banks until these marginal are equalized. If instead the marginal value of private consumption in a bank is lower than the marginal value of public consumption, the bank will not be bailed out and the constraint in (2.2) will bind.

The following two propositions characterize the key features of the constrained efficient allocation of resources in our environment. First, the consumption of investors in banks with sound fundamentals is independent of the aggregate state and these banks do not receive bailouts.[6]

Proposition 2.1. *The constrained efficient allocation satisfies*

$$(c_{10}^*, c_{20}^*) = (c_{1S}^*, c_{2S}^*) \quad \textit{and} \quad b_S^* = 0.$$

Given this result, we will drop the (c_{10}, c_{20}) notation in what follows and use (c_{1S}, c_{2S}) to refer to the consumption profile for investors in a bank with sound fundamentals regardless of the aggregate state. Our second result shows that this profile is different from the one assigned by the planner to investors in banks with weak fundamentals.

[6]The first part of this result depends on our simplifying assumption that sound banks are completely unaffected by the bad aggregate state, but the second part of the result does not. Even if sound banks were to experience some losses during an economic downturn, the planner would not choose to bail out these banks as long as the losses are small relative to those at weak banks.

Proposition 2.2. *The constrained efficient allocation satisfies*

$$(c_{1S}^*, c_{2S}^*) \gg (c_{1W}^*, c_{2W}^*) \quad \text{and} \quad b_W^* > 0.$$

This result shows that the constrained efficient involves a combination of *bailouts* and *bail-ins* for investors in banks with weak fundamentals. The optimal bailout b_W^* gives investors partial insurance against the risk associated with their bank's losses. However, the consumption of investors in weak banks remains below that of investors in sound banks; this difference can be interpreted as the degree to which the planner "bails in" the investors in weak banks. The efficient level of insurance is only partial in this environment because offering insurance is costly; it requires the planner to collect more tax revenue, which leads to an inefficiently high level of the public good in the good aggregate state.

It is worth pointing out that the constrained-efficient bail-in applies equally to *all* investors in a weak bank, regardless of when they arrive to withdraw. While the desirability of this feature follows immediately from risk aversion, we will see below that it often fails to hold in a decentralized equilibrium. It is also worth noting that the constrained efficient allocation is incentive compatible. The first-order conditions (2.3) and $R > 1$ imply that $c_{1z}^* < c_{2z}^*$ holds for every state z and, hence, a patient investor always prefers her allocation to that given to an impatient investor (and vice versa).

2.4 Equilibrium within a bank

In this section we begin our investigation of the decentralized economy. Compared to the planner's economy discussed in the previous section, the decentralized economy is different in the following important ways. First, investors' preference types are private information and banks therefore allow investors to choose in which period they withdraw. Second, each bank is concerned solely with its own investors and takes economy-wide variables, including the level of the public good, as given. Third, there is asymmetric information between the banks and the government; while the government immediately observes the aggregate state at the beginning of $t = 1$, it must wait for θ withdrawals to take place before observing bank-specific states. The government then makes bailout payments to banks with weak fundamentals and places these bank into resolution. Importantly, the bailout and resolution policies cannot be set ex-ante, but instead are chosen as a best response to the situation at hand.

In this section, we study equilibrium in the withdrawal game played by an individual bank's investors, taking the actions of investors at other banks (and the government) as given. In section 2.5, we study the joint determination of equilibrium actions across all banks.

2.4.1 Preliminaries

We begin by reviewing the timeline of events in Figure 2.1 for the decentralized economy and then provide a general definition of equilibrium.

The tax rate. To simplify the analysis in this section, we assume that the tax rate τ levied by the government in period 0 is set to the value from the constrained efficient allocation, τ^*. We derive equilibrium withdrawal behavior and the equilibrium allocation of resources for this tax rate.

Banking contracts. In period 0, each bank establishes a contract that specifies how much it will pay to each withdrawing investor as a function of both the bank's fundamental state $\sigma_k \in \{0, \bar{\sigma}\}$ and the fraction $\rho^k \in [\pi, 1]$ of its investors who choose to withdraw early. We allow the government to set an upper bound $\bar{c}$ on the payments made to any investor withdrawing in period 1. One way to justify this upper bound is to assume that while the government cannot dictate the exact terms of the contractual arrangement between a bank and its investors, it is able to impose broad guidelines on the types of contract banks are allowed to offer.

Because investors are risk averse, it will be optimal for a bank to give the same level of consumption to all investors who withdraw in the same period.[7] The operation of the bank is, therefore, completely described by the function

$$c_1^k : \{0, \bar{\sigma}\} \times [\pi, 1] \rightarrow [0, \bar{c}], \tag{2.6}$$

where c_1^k denotes the payment made by the bank to each investor who withdraws in period 1. The bank's matured investment in period 2, plus any bailout payment received, is divided evenly among its remaining depositors. We refer to the function in (2.6) as the *banking contract.* There is full commitment to the banking contract in the sense that the plan in (2.6) will be followed unless the bank is placed into resolution by the government. Each bank's contract is chosen to maximize the expected utility from private consumption of the bank's investors.[8]

Bailouts and resolution. After a fraction θ of investors have withdrawn at $t = 1$, the government observes the fundamental state σ_k of each bank and chooses a bailout payment b^k for each bank with weak fundamentals. It then dictates the payments made by these banks to their remaining investors as part of the resolution process. We characterize the government's bailout/resolution policy below.

Withdrawal strategies. An investor's withdrawal decision can depend on both her preference type ω_i^k and the fundamental state of her bank σ_k. (See Figure 1.) A withdrawal strategy for investor i in bank k is, therefore, a mapping:

[7] Keep in mind that our environment is different from that studied by Wallace (1990), Green and Lin (2003), Peck and Shell (2003) and others where the bank gradually learns about the demand for early withdrawal by observing investors' actions as they take place. Here, a bank directly observes total early withdrawal demand before making any payments to investors. It learns no new information as investors sequentially withdraw at $t = 1$ and, therefore, an optimal arrangement will always give the same level of consumption to each of these investors.

[8] This outcome would obtain, for example, if multiple banks competed for deposits in each location. We use a representative bank in each location only to simplify the presentation.

$$y_k^i : \Omega \times \Sigma \to \{0, 1\} \tag{2.7}$$

where $y_k^i = 0$ corresponds to withdrawing in period 1 and $y_k^i = 1$ corresponds to withdrawing in period 2. An investor will always choose to withdraw in period 1 if she is impatient. We introduce the following labels to describe the actions an investor takes in the event she is patient.

Definition 2.3. For given σ_k, we say investor i in bank k follows:

(i) the *no-run strategy* if $y_k^i\left(\omega_k^i, \, \sigma_k\right) = \omega_k^i$ for $\omega_k^i \in \{0, 1\}$, and

(ii) the *run strategy* if $y_k^i\left(\omega_k^i, \, \sigma_k\right) = 0$ for $\omega_k^i \in \{0, 1\}$.

We use y_k to denote the profile of withdrawal strategies for all investors in bank k and y to denote the withdrawal strategies of all investors in the economy. It will often be useful to summarize a profile of withdrawal strategies by the fraction of investors who follow the run strategy in that profile, which we denote

$$x_{\sigma_k} \equiv \int\limits_{i \in [0,1]} (1 - y_k^i(\omega_k^i = 1, \sigma_k))dk.$$

Similarly, we use ρ_k to denote the total demand for early withdrawal from bank k in a given profile, which equals

$$\rho_k = \pi + (1 - \pi)x_k.$$

Allocations. The allocation of private consumption in bank k in a particular state is a specification of how many investors withdraw at $t = 1$ in that state, how much consumption each of these investors receives, and how much consumption each remaining investor receives at $t = 2$. This allocation depends on the banking contract for bank k, the withdrawal strategies of investors in bank k, and the government intervention in bank k (if any). The details of the government intervention, in turn, may depend on the contracts of other banks and the withdrawal strategies of investors in those banks. In general, therefore, the optimal withdrawal behavior for each investor in bank k may depend on the contracts offered by other banks on and the withdrawal strategies of investors in other banks.

Equilibrium. To study equilibrium withdrawal behavior within a single bank, we fix all banking contracts, the government's intervention policy, and the withdrawal strategies of investors in all other banks, y_{-k}. Together, these items determine the payoffs of what we call the *withdrawal game in bank k*. That is, holding these other items fixed, we can calculate the allocation of private consumption in bank k as a function of the strategies y_k played by that bank's investors. An equilibrium of this game is a profile of strategies for the bank's investors, y_k, such that for every investor i in the bank, y_k^i is a best response to the strategies of the other investors, y_k^{-i}.

An equilibrium of the overall economy is a profile of withdrawal strategies for all investors y^* such that (i) y_k^* is an equilibrium of the withdrawal game in bank k generated by the strategies y_{-k}^* of investors in all other banks, for all k, (ii) the contract in bank k maximizes the expected utility of its investors taking as given the contracts and withdrawal strategies y_{-k}^* of investors in all other banks, and (iii) the government's bailout and resolution policy maximizes total welfare taking as given all banking contracts and withdrawal strategies y^*. Notice how this definition reflects the timing assumptions depicted in Figure 1. Investors in bank k recognize that their choice of contract will influence equilibrium withdrawal behavior within their own bank but will not affect outcomes at other banks.[9] The government's bailout and resolution policies, in contrast, are set after all banking contracts and withdrawal decisions have been made. Because the government cannot pre-commit to the details of these policies, it acts to maximize welfare taking all bank contracts and withdrawal

[9]This result follows, in part, from the assumption that there are a continuum of locations and, hence, the actions taken at one bank have no effect on aggregate variables or on the behavior of the government toward other banks.

decisions as given.

In the subsections that follow, we derive the properties of the contracts that a bank will use in equilibrium, focusing first on the case where its fundamental state is strong. We then turn to the case where the bank's fundamental state is weak, which requires characterizing the optimal bailout and resolution policies as well.

2.4.2 Banks with sound fundamentals

We assume the government does not give bailouts to banks with sound fundamentals, nor does it place them in resolution.[10] As a result, all investors who chose to withdraw at $t = 1$ receive the contractual amount $c_1^k(0,\, \rho^k)$, and all investors who chose with withdraw at $t = 2$ receive an even share of the bank's assets, $c_2^k(0,\, \rho^k)$, which is implicitly defined by

$$p^k c_1(0,\, \rho^k) + (1 - \rho^k)\frac{c_2(0,\, \rho^k)}{R} = 1 - \tau. \tag{2.8}$$

The bank and its investors recognize that ρ^k will result from the equilibrium withdrawal behavior of investors. In particular, if the bank offers a higher payment in period 1 than in period 2, all investors will choose to withdraw early. In other words, equilibrium requires

$$\rho^k = \left\{ \begin{array}{c} \pi \\ \in [\pi, 1] \\ 1 \end{array} \right\} \quad \text{as} \quad c_1(0,\, \rho^k) \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} c_2(0,\, \rho^k). \tag{2.9}$$

We refer to (2.9) as the *implementability constraint.* If a triple (ρ_S, c_{1S}, c_{2S}) satisfy both (2.8) and (2.9), then any banking contract with $c_1^k(\sigma_k, \rho_S) = c_{1S}$ for $\sigma_k = 0$ can implement this allocation as an equilibrium of the withdrawal game in bank k,

[10] Recall from Section 2.3 that the constrained efficient allocation involves zero bailouts for sound banks. Our assumption here is that the government is able to commit to follow this policy.

regardless of how the payments $c_1^k(\sigma_k, \rho^k)$ are set for other values of ρ^k. The following result shows that something stronger is true: by choosing these other payments appropriately, the banking contract can be set so that the withdrawal game in bank k has a *unique* equilibrium.

Proposition 2.4. *If the allocation* (ρ_S, c_{1S}, c_{2S}) *satisfies both* (2.8) *and* (2.9)*, then there exist a contract which implements this allocation as the unique equilibrium of the withdrawal game played by bank k's investors conditional on the bank being sound.*

In light of the above proposition, we can recast the problem of choosing the optimal banking contract as one of directly choosing the allocation (ρ_S, c_{1S}, c_{2S}) to maximize expected utility

$$V_S\left(\rho_S,\, c_{1S}\right) \equiv \rho_S u(c_{1S}) + (1-\rho_S)u(c_{2S}) \tag{2.10}$$

subject to the feasibility constraint (2.8) and the implementability constraint (2.9) and the restriction $c_1^k \in [0, \bar{c}]$ for all banks. The next result characterizes the solution to this problem.

Proposition 2.5. *When bank k has sound fundamentals, there is a unique equilibrium of the withdrawal game in bank k associated with the optimal banking contract. The equilibrium allocation* (ρ_S, c_{1S}, c_{2S}) *satisfies* $\rho_S = \pi$ *and* $c_{1S} = min\{c_{1S}^*, \bar{c}\}$*.*

This result shows that as long as the upper bound $\bar{c}$ is set high enough to allow it, the equilibrium allocation within a sound bank is the same as in the constrained efficient allocation. In other words, in the absence of government intervention, banks will allocate resources efficiently in our model and the problem of bank runs will not arise. One contract that would uniquely implement the desired allocation is

$$c_1^k\left(0,\, \rho^k\right) = \left\{ \begin{array}{c} min\{c_{1S}^*, \bar{c}\} \\ 0 \end{array} \right\} \quad \text{if} \quad \left\{ \begin{array}{c} \rho^k = \pi \\ \rho^k > \pi \end{array} \right\} \tag{2.11}$$

If only impatient investors withdraw in period 1, we have $\rho^k = \pi$ and therefore the contract in (2.11) sets $c_1^k\left(0,\,\rho^k\right)$ to the lower of c_{1S}^* and $\bar{c}$. The feasibility constraint in (2.8) then implies that $c_1^k\left(0,\,\rho^k\right) < c_2^k\left(0,\,\rho^k\right)$ holds and, therefore, patient investors strictly prefer to withdraw in period 2. On the other hand, if a positive measure of the patient (in addition to the impatient) investors withdraw in period 1, then $\rho^k > \pi$. The contract in (2.11) would then set $c_1^k\left(0,\,\rho^k\right) = 0$, which is a (strong) form of suspending withdrawals when faced with a run. Given this payoff, and $c_2^k\left(0,\,\rho^k\right)$ from the budget constraint in (2.8), withdrawing early is clearly not a best response for any of the bank's patient investors. Therefore, with a contract as in (2.11), there is a unique allocation consistent with equilibrium.

2.4.3 Banks with weak fundamentals

In the bad aggregate state, a fraction $n > 0$ of the banks will have weak fundamentals. Let W denote the set of weak banks,

$$W \equiv \{k \in [0,1] \text{ s.t } \sigma_k = \bar{\sigma}\} \tag{2.12}$$

After the first θ withdrawals have taken place in all banks, the government observes the fundamental state σ_k of each bank. For banks with weak fundamentals, the government also observes the bank's current condition: the amount of resources remaining in the bank and the fraction of the bank's remaining investors who are impatient. The government then decides on a bailout payment b^k for each $k \in W$ and places these banks into resolution.

To simplify the presentation, we assume that when a bank is placed into resolution, the government directly observes the preference types of the bank's remaining investors and allocates the bank's resources (including the bailout payment) conditional on these types.[11] This assumption is **not** important for our results. If instead the government were to offer a new banking contract and have the remaining investors play a withdrawal game based on this new contract, it could choose a contract that yields the outcome we study here as the unique equilibrium of that game.

Resolution. Let $\hat{\psi}_k$ denote the per-capita level of resources in bank k, including any bailout payment received, after the first θ withdrawals have taken place. That is,

$$\hat{\psi}_k = \frac{(1-\tau)(1-\bar{\sigma}) - \theta c_1^k\left(\bar{\sigma},\, \rho_k\right) + b_k}{1-\theta} \tag{2.13}$$

where b_k is the per-investor bailout given to bank k. Let $\hat{\rho}^k$ denote the fraction of the bank's remaining investors who are impatient. The allocation of resources for a bank in resolution is chosen to maximize the sum of the utilities for the remaining investors in the bank:

$$\hat{V}\left(\hat{\psi}_k;\, \hat{\rho}_k\right) \equiv \underset{\hat{c}_1^k,\, \hat{c}_2^k}{max}\ (1-\theta)\left(\hat{\rho}_k u\left(\hat{c}_1^k\right) + (1-\hat{\rho}_k) u\left(\hat{c}_2^k\right)\right) \tag{2.14}$$

subject to the feasibility constraint

$$\hat{\rho}_k \hat{c}_1^k + (1-\hat{\rho}_k)\frac{\hat{c}_2^k}{R} \leq \hat{\psi}_k \tag{2.15}$$

The optimal choice of post-bailout payments is determined by the first order condition

$$u'\left(\hat{c}_1^k\right) = Ru'\left(\hat{c}_2^k\right) = \hat{\mu}(\hat{\psi}_k;\, \hat{\rho}_k), \tag{2.16}$$

where $\hat{\mu}$ is the Lagrange multiplier on the resource constraint. Since $R > 1$, this condition implies that a bank under resolution provides more consumption to patient investors withdrawing in period 2 than to the remaining impatient investors who withdraw in period 1.

[11] One could imagine, for example, the using the court system to evaluate individual's true liquidity needs, as discussed in Ennis and Keister (2009).

Bailouts. In choosing the bailout payments $\left\{b^k\right\}$, the government's objective is to maximize the sum of the utilities of all investors in the economy. Given that bailout payments are only made to banks with weak fundamentals, this problem can be written as

$$\max_{\left\{b^k\right\}_{k\in W}} \int_W \hat{V}\left(\hat{\psi}_k;\, \hat{\rho}_k\right) dk + v\left(\tau - \int_W \hat{b}_k dk\right) \tag{2.17}$$

Notice that while bailout payments raise the private consumption of investors in weak banks, they lower the provision of the public good, which affects all investors. The first-order condition for this problem can be written as

$$\hat{\mu}(\hat{\psi}_k;\, \hat{\rho}_k) = v'\left(\tau - \int_W \hat{b}_k dk\right) \text{ for all } k. \tag{2.18}$$

Notice that the right-hand side of this equation – the marginal utility of public consumption – is independent of k. The optimal bailout policy thus has the feature that the marginal value of resources will be equalized across all weak banks, regardless of their chosen banking contract or the withdrawal behavior of their investors. Let $(\hat{c}_1, \hat{c}_2)$ denote the common consumption allocation that will be given to impatient and patient investors, respectively, in banks that are under resolution. Then the bailout payment b^k made to bank k will be chosen so that the following feasibility constraint is satisfied

$$\hat{\rho}_k\hat{c}_1 + (1-\hat{\rho}_k)\frac{\hat{c}_2}{R} = \frac{(1-\tau)(1-\bar{\sigma}) - \theta c_1^k\left(\bar{\sigma},\, \rho_k\right) + \hat{b}_k}{1-\theta}. \tag{2.19}$$

Observe that the bailout payment made to bank k is increasing in the amount paid out by the bank before being bailed out, $c_1^k\left(\bar{\sigma},\, \rho_k\right)$, and in the fraction of investors its remaining investors who are impatient, $\hat{\rho}_k$.

Banking contract. In the event that the bank's fundamentals are weak, its investors recognize that a fraction θ of them will receive the amount specified by the contract, $c_1^k(\bar{\sigma},\, \rho_k)$, before the government intervenes. At this point, the bank will be bailed out and placed into resolution. Its remaining impatient investors will receive $\hat{c}_1$ and its remaining patient investors will receive $\hat{c}_2$, as derived above. In choosing the contract for this case, the bank recognizes that its investors will choose to run if it sets $c_1^k > \hat{c}_2$ and will not run if $c_1^k < \hat{c}_2$. In other words, the fraction ρ^k of investors who attempt to withdraw at $t = 1$ will satisfy

$$\rho^k = \left\{ \begin{array}{c} \pi \\ \in [\pi, 1] \\ 1 \end{array} \right\} \text{ if } c_{1W} \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} \hat{c}_2. \tag{2.20}$$

Proposition 2.6. *If* $(\rho_W,\, c_{1W})$ *satisfy* (2.20)*, then there exists a banking contract* c_1^k *that implements this allocation as the unique equilibrium of the withdrawal game played by bank k's investors when the bank is weak.*

As in the case of sound banks studied above, we can equivalently formulate bank's problem of choosing the optimal contract as one of directly choosing the allocation $(\rho_W,\, c_{1W})$ to maximize

$$V_W(\rho_W,\, c_{1W}) \equiv \theta u(c_{1W}) + (1-\theta)\left[\hat{\rho}_{\bar{\sigma}} u(\hat{c}_1) + (1-\hat{\rho}_{\bar{\sigma}}) u(\hat{c}_2)\right] \tag{2.21}$$

subject to the implementability constraint for weak banks (2.20) and the relationship

$$\hat{\rho}_{\bar{\sigma}} \equiv \frac{\pi}{1-\theta}\left(\frac{\rho_W - \theta}{\rho_W}\right). \tag{2.22}$$

This last expression shows how the fraction of the bank's remaining investors after θ withdrawals are impatient depends on the fraction that initially attempt to withdraw early. The first term in the objective function in (2.21) is clearly increasing in the choice of c_{1W}. However, the implementability constraint (2.20) shows that if c_{1W} is set greater than $\hat{c}_2$, the bank's investors will run in which case ρ^k will be equal to 1. A run is costly for the bank's investors because $\hat{\rho}_{\bar{\sigma}}$ is an increasing function of ρ^k and the second term in the objective function (2.21) is strictly decreasing in $\hat{\rho}_{\bar{\sigma}}$.

Proposition 2.7. *The solution to the program of maximizing* (2.21) *subject to* (2.20) *and* (2.22) *will either set* $c_{1W} = \bar{c}$ *or* $c_{1W} = \hat{c}_2$.

If $c_{1W} \leq \hat{c}_2$ then there is no run on the bank and the sum of utilities for its investors equal $V_W(c_{1W}, \pi)$. On the other hand, setting $c_{1W} > \hat{c}_2$ leads to a run and the sum of utilities for its investors would equal $V_W(c_{1W}, 1)$. A weak bank can strictly gain by setting c_{1W} equal to the upper bound $\bar{c}$ whenever

$$V_W(\bar{c}, 1) > V_W(\hat{c}_2, \pi) \tag{2.23}$$

The above inequality implies that the loss to the remaining $1-\theta$ investors in the bank resulting from keeping payments as high as possible is more than offset by the gain to the first fraction θ to withdraw, that is, a weak bank will have no incentive to lower its payment to $\hat{c}_2$, even if this would prevent its investors from running. Next, by observing that the inequality in (2.23) is equivalent to $u(\bar{c}) - u(\hat{c}_2) > (1-\pi)(u(\hat{c}_2) - u(\hat{c}_1))$, the equilibrium outcomes in weak banks can be characterized as in the following proposition:

Proposition 2.8. *If bank k has weak fundamentals then:*

(i) If $u(\bar{c}) - u(\hat{c}_2) < (1-\pi)(u(\hat{c}_2) - u(\hat{c}_1))$, *there is a unique equilibrium of the withdrawal game in bank k associated with the optimal banking contract. The equilibrium allocation has* $\rho_W = \pi$ *and* $c_{1W} = \min\{\bar{c}, \hat{c}_2\}$.

(ii) If $u(\bar{c}) - u(\hat{c}_2) > (1-\pi)(u(\hat{c}_2) - u(\hat{c}_1))$, *there is again a unique equilibrium of the withdrawal game in bank k associated with the optimal banking contract. The equilibrium allocation in this case has* $\rho_W = 1$ *and* $c_{1W} = \bar{c}$.

(iii) If $u(\bar{c}) - u(\hat{c}_2) = (1-\pi)(u(\hat{c}_2) - u(\hat{c}_1))$, *the withdrawal game in bank k has*

multiple equilibria, one with $\rho_W = \pi$ *and* $c_{1W} = \hat{c}_2$ *and another with* $\rho_W = 1$ *and* $c_{1W} = \bar{c}$.

Among other things, Proposition 2.8 shows that, in some cases, weak banks experience a run as part of equilibrium, even though they learn right away that a run is under way and payments can be made fully contingent on both the bank's fundamental and the demand for early withdrawals. If $V_W(\hat{c}_2, \pi) < V_W(\bar{c}, 1)$ then one optimal contract for bank k is the following:

$$c_1^k\left(\bar{\sigma}, \rho^k\right) = \bar{c} \quad \text{for} \quad \rho^k \in [\pi, 1]\,. \tag{2.24}$$

Under this contract, bank k always pays the maximal amount $\bar{c}$ to the first θ of its investors to withdraw, even if this leads to a run. On the other hand, if If $V_W(\hat{c}_2, \pi) > V_W(\bar{c}, 1)$ then then an optimal contract for bank k will be

$$c_1^k\left(\bar{\sigma}, \rho^k\right) = \left\{ \begin{array}{c} \min\{\bar{c}, \hat{c}_2\} \\ \\ 0 \end{array} \right\} \quad \text{for} \quad \left\{ \begin{array}{c} \rho^k = \pi \\ \\ \rho^k > \pi \end{array} \right\}. \tag{2.25}$$

This second contract ensures that, if weak, bank k sets its payment c_{1W} as high as possible, but not above $\hat{c}_2$. In this case there is no run and only impatient investors withdraw in period 1.

2.5 Equilibrium across banks

The previous section investigated the equilibrium outcomes within a given bank, taking the actions of the government and the remaining banks as fixed. We now investigate the properties of the overall equilibrium across all banks.

2.5.1 Constrained inefficiency

We begin by asking whether the equilibrium allocation is constrained efficient. Note that, in order for this allocation to be feasible in the decentralized economy, the upper bound $\bar{c}$ on early payments must be set sufficiently high that sound banks are able to choose c^*_{1S}. For the analysis in this section, we will set $\bar{c} = c^*_{1S}$. Our next result shows that, even though it is feasible, the constrained efficient allocation is never an equilibrium of the decentralized economy.

Proposition 2.9. *The equilibrium allocation of resources is never constrained efficient.*

The bailout policy creates an incentive for weak banks to set their early payments as high as possible. The only reason a weak bank would voluntarily impose losses on its investors (by setting a payment below $\bar{c} = c^*_{1S}$) is to prevent a run. Note that preventing a run only requires that the payment in period 1 not exceed $\hat{c}_2$ and, as a result, a weak bank will never set its early payment below this level. In particular, a weak bank will never choose to bail in its investors all the way down to $\hat{c}_1$, as occurs in the constrained efficient allocation.

2.5.2 Equilibrium bank runs

In addition to being constrained inefficient, the equilibrium of the full model will, in some cases, involve a run by investors on weak banks.

Proposition 2.10. *For some parameter values, there exists an equilibrium in which investors run on weak banks. In some cases this equilibrium is unique, but in others it coexists with another equilibrium in which no run occurs.*

In the run equilibrium, all investors in weak banks attempt to withdraw at $t = 1$, that is, the profile of withdrawal strategies has $x_{\bar{\sigma}} = 1$. A fraction θ of these investors successfully withdraw before the government observes $\sigma_k = \bar{\sigma}$ and places the bank into resolution. The result in Proposition 2.10 is established on Figure 2.2 which depicts the type of equilibria that arise for different combinations of the parameters n, the fraction of weak banks, and $\bar{\sigma}$, the amount lost by each of them. The figure uses the utility function[12]

$$u(c_1^{i,k} + \omega^{i,k} c_2^{i,k}) = \frac{\left(c_1^{i,k} + \omega^{i,k} c_2^{i,k}\right)^{1-\gamma} - 1}{1-\gamma} \qquad \text{and} \qquad v(g) = \delta \frac{g^{1-\gamma}}{1-\gamma}. \tag{2.26}$$

For parameter combinations in the dark region in the lower-left part of the graph, there is a unique equilibrium of the model and the allocation in this equilibrium does not involve a bank run. When the losses $\bar{\sigma}$ suffered by a weak bank are small and/or few banks experience these losses (i.e., n is small), the process of resolving these banks has a relatively small cost for the government. When this cost is small, the government remains in good fiscal condition and will choose to make bailout payments that lead to relatively high consumption levels $(\hat{c}_1, \hat{c}_2)$ for the remaining investors in banks placed into resolution. This fact, in turn, makes running in an attempt to withdraw before the government intervenes less attractive for patient investors in a weak bank. As a result, a unique equilibrium exists and all patient investors wait until $t = 2$ to withdraw.

In the unshaded region in the upper-right portion of the figure, in contrast, both the number of banks experiencing a loss and the amount lost by each of these banks are significant. In this case, the government's budget constraint will be substantially impacted by its desire to bail out weak banks in a crisis. As the marginal value of public resources rises, the bailout and resolution process will lead to lower consumption levels $(\hat{c}_1, \hat{c}_2)$ for the remaining investors in these banks. When $\hat{c}_2$ is low

[12] The other parameters of the model are set to $R = 1.5$, $\pi = 0.5$, $\gamma = 5$, $\delta = 0.5$, $q = 0.05$ and $\theta = 0.5$. The tax rate τ is set to its constrained efficient value from section 2.3.

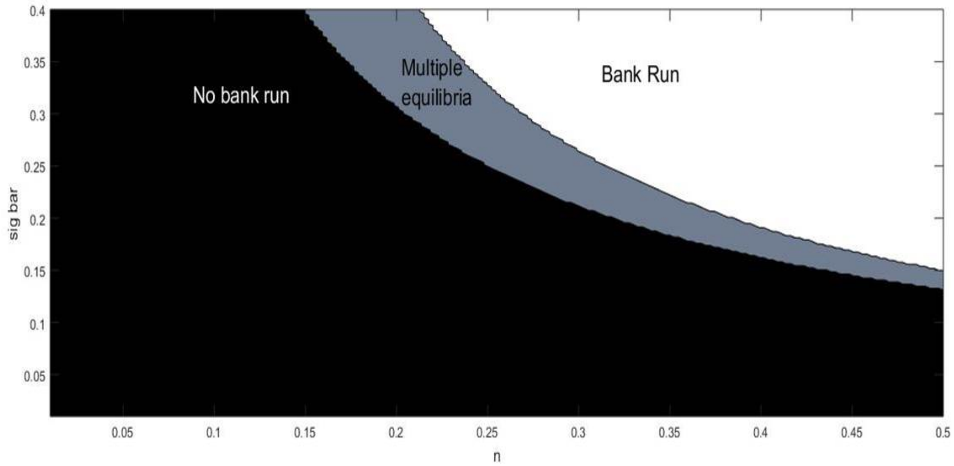


Figure 2.2: Equilibrium with a bank run

enough, the equilibrium within a weak bank k will involve a run by patient investors, as shown in Proposition 2.8. The overall equilibrium in this region is still unique, but the (larger) losses on weak banks' asset are now compounded by the additional liquidation of assets and misallocation of resources created by the run.

In between these two regions, both of the equilibria described above exist. The fact that multiple equilibria exist in this region is particularly interesting in light of Proposition 2.8, which showed that the equilibrium of the withdrawal game within each bank is unique expect for in a knife-edge case. The multiplicity of equilibria illustrated in Figure 2.2 arises because of an externality in payoffs *across* weak banks. When a run occurs at other weak banks, this event causes more investment to be liquidated and leads to larger bailouts at those banks. The larger bailouts place greater strain on the government's budget constraint and lead – all else being equal – to a smaller bailout at bank k. In the lighter-shaded region in Figure 2.2, this smaller bailout lowers the consumption levels $(\hat{c}_1, \hat{c}_2)$ enough to make running a best response for the patient investors in bank k. In other words, in our model there is a strategic complementarity in the withdrawal decisions of investors across banks. The usual strategic complementarity that appears in models in the Diamond-Dybvig tradition – which arises between investors within a bank – is eliminated by the more flexible banking contracts. However, the government's bailout and resolution policy introduces this new complementarity in actions across banks, which creates the region of multiple equilibria in Figure 2.2.

It is worth emphasizing that a run on bank k lowers the welfare of the bank's investors in much the same way as in the existing literature. Holding fixed the bailout payment it receives, a bank's investors would be strictly better off if there were no run on the bank. Moreover, the bank has contractual tools that would allow it to prevent the run. The problem, however, is that preventing the run requires decreasing the payment given to the first θinvestors who withdraw and this action would decrease the bailout payment the bank receives. Instead, in this equilibrium, the bank's investors choose to tolerate the run as a side effect of the plan that maximizes the level of payments the bank is able to make to its investors before the government intervenes.

2.5.3 The importance of real losses

Proposition 2.5 established that patient investors never run on a sound bank in equilibrium. In this sense, real losses at weak banks are a necessary ingredient for a bank run to occur in our model. The patterns in Figure 2.2 suggest that, in addition, the bank run equilibrium tends to exist when the total losses in the bad aggregate state are large and tends not to exist when these losses are small. The following result shows one sense in which this result holds more generally.

Proposition 2.11. *Given other parameter values, there exists $\bar{n} > 0$ such that for all $n < \bar{n}$, there is no bank run in equilibrium.*

Unlike the traditional Diamond-Dybvig model, where a run can occur on a single bank, a run in our model cannot be an isolated event. If the number of affected banks is small, the associated losses will have a minimal effect on the government's budget constraint. If the government remains in good fiscal condition, the bailout policy it will choose ex post treats weak banks generously, leaving their patient investors with no incentive to run.

In summary, a bank run in our environment cannot occur unless (i) banks sustain real losses, (ii) these losses are reasonably wide-spread, and (iii) the government is slow to react. In other words, a run here is necessarily systemic, triggered by a real shock and accompanied by a sluggish policy response to the unfolding crisis.

Discussion

In this section we relate our results to a number of recent policy proposals that allow for more state contingent contracts with the goal of improving intermediaries' capacity to deal with periods of distress. We argue that contractual arrangements designed to promote financial stability such as various "bail-in rules" are unlikely to be sufficient to eliminate the problem of bank runs in an environment characterized by limited commitment, asymmetric information and bailouts.

Bail-in options to promote stability. Bank runs in our environment are not based on agency costs – this observation is important and deserves further emphasis. A number of recent legislative changes aim to promote financial stability by endowing financial intermediaries with increased contractual flexibility, which would allow them to react as soon as they start to experience distress. For example, "gates" and withdrawal fees in money market mutual funds, swing pricing in the mutual fund industry more generally, and the new bail-in rules in the US, Europe and elsewhere can all be interpreted as giving intermediaries the opportunity – but not necessarily the obligation – of imposing losses on all (or subset) of their investors if this is deemed desirable for the long term health of the institution. The hope of these legislative reforms is that these new "bail-in options" would not only be effective in mitigating fragility (or even preventing runs entirely), but in addition, would eliminate the need for taxpayers to finance a bailout or at least drastically reduce the cost of government's interventions.

For instance, the objective of the recent amendments to *Rule 2a-7* which allows money market mutual funds in the U.S. to impose withdrawal fees and "gates" under certain conditions can be readily interpreted in our framework. Indeed, the purpose of these new tools is to reduce investors' incentive to redeem quickly and ahead of others (i.e. to run) when the fund is in distress. At the same time, the imposition of fees and gates must be subject to the approval of the board of directors, who have the discretion to use these tools only if this is determined to be in the best interest of their shareholders. Notice that from the perspective of our model, withdrawal fees and "gates" can be captured as setting lower payments in weak banks. In this case, our results show that if bailouts are anticipated, then then these "bail-in options" would be underused or even avoided entirely, and thus fail to promote financial stability.

Bailouts "crowd-out" bail-ins. An effective regulatory framework must incorporate the fact that bailouts have the tendency to "*crowd-out*" bail-ins. One way of addressing this issue is to prevent intermediaries from avoiding or delaying the imposition of bail-in on their investors. A given intermediary might have discretion in choosing the size and timing of the bail-in because (i) the regulatory framework does not explicitly precludes them from doing so or (ii) the government may lack information or expertise to effectively impose bail-in rules.

Indeed, the stated purpose of a large part of the new "bail-in" legislation is to provide financial institutions with the discretion to use these tools while operating in the best interest of their stakeholders. As we have seen, however, operating in the best interest of their own stakeholders leads to inferior social outcomes if banks anticipate being bailed out by the government. The reason is that individual banks do not internalize the extra cost imposed on taxpayers by their delayed reactions.

At the same time, ensuring that banks adhere to strict bail-in rules requires the imposition of penalties for failing to comply. These penalties, however, will end up being fully passed onto the investors in weak banks and thus depress payments even further exactly at a time these banks were already suffering losses. As a result - in a form of regulatory forbearance - a benevolent policy maker will be ex-post unwilling to impose additional cost on weak banks.[13] [14]

Delays in the policy maker's intervention. The bailout intervention takes places after a fraction θ of the investors in weak banks have already withdrawn. As discussed above, this delayed reaction from the policy maker is one of the primary ingredients creating financial fragility in our setting. This slow response can occur for variety of reasons. In the benchmark model, this slow reaction was based on informational frictions – the policy maker was unaware initially which banks had suffered losses and obtains this information only when the measure of withdrawals reaches θ. There, are other ways to motivate such a delay. For example, θ can be interpreted more generally as a sum of two components $\theta = E + R$, where E is the time necessary to carry examinations and learn the subset of weak banks and R is the time necessary to implement the resolution mechanism. Effective supervision is thus associated with lower E, whereas regulatory systems - such as an orderly resolution authority (OLA) – with a lower R.

The slow reaction of the government can also be interpreted as reflecting political power from certain special interest groups. For example, investors who are well-connected politically might also be among the first with an opportunity to withdraw during times of financial distress. These investors could then use their political influence to convince the policy maker not to step in right away and instead to impose the

[13]In addition, establishing that a given financial intermediary knowingly deviated from the bail-in rules set up by the government could be far from straightforward, especially if a court system places the burden of proof on the regulator.

[14]A strict no-bailout rule eliminates banks runs in our model, but may be difficult to credibly commit to. Moreover, even if the government were able to commit not to provide bailouts, doing so would **not** necessarily improve welfare. The reason is that bailouts provide socially valuable ex-post consumption insurance for the investors in weak banks. In fact, we can easily find parameter values for which a strict no-bailout rule would reduce ex-ante welfare, in some cases to a significant extent (i.e. when $\bar{\sigma}$ is large).

haircuts on those withdrawing later on. In addition, the timing of the government intervention might reflect opaque incentives faced by regulators.[15]

Another reason for a delay in the banking resolution might be related to political timing. Brown and Dinc (2005) study episodes of government resolutions of failed banking institutions in 21 major emerging markets during the 1990s and provide evidence that the timing of a government's intervention depends on the electoral cycle. In particular, costly government interventions which would impose a high cost on the taxpayer and would also fully reveal the extent of the financial crisis (and thus may raise questions of how the government allowed this to happen in the first place) were significantly less likely to occur before elections. In other words, political factors might lead inaction or to a delay in the adoption of beneficial economic policies. (See also Rogoff and Sibert, 1988.)

2.6 Conclusion

A necessary ingredient for a bank run to occur in the the bank in question be slow to react to the surge in withdrawal demand. This slow reaction is what leads investors to anticipate that the future payments made by the bank will be smaller and, hence, gives them an incentive to try to withdraw before the reaction comes. In the previous literature, the primary factors behind this slow response have been exogenously imposed rather than derived endogenously as part of the equilibrium outcome. Specifically, banks' failure to respond in a timely manner has been justified by assuming that either (i) contracts are rigid and therefore cannot be ex-post altered to deal with

[15]Kroszner and Strahan (1996) argue that throughout the eighties the Federal Savings and Loan Insurance Corporation (FSLIC) was faced with a severe shortage of cash with which to resolve insolvent thrift institutions. This lack of funds forced the FSLIC to practice regulatory forbearance and to delay its explicit intervention in insolvent mutual thrifts in anticipation that the government would eventually supply additional resources. This delay led a large number of insolvent thrift institutions to maximize the value of future government liabilities guarantees (at the taxpayers' expense) by continuing to pay high dividends until the eventual resolution mechanism was put in place.

a run, or (ii) banks were unable to respond efficiently to a run because they were (at least initially) unaware that the run was actually taking place.

In contrast, we have presented a model of banking and government interventions where (i) banks maximize the utilities of their investors (i.e. there are no agency costs), (ii) contracts can be made fully state contingent, and (iii) banks always have sufficient information to respond in a timely and effective way to an incipient run. In common with the existing literature, a bank run in our setting can occur only when the bank's reaction to the run is delayed. However, the delayed reaction in our model is the endogenous choice of the bank, acting in the best interests of its investors. We show that this framework has a number of interesting implications. For example, banks will not have an incentive to use bail-in options and similar measures to impose discretionary losses on their investors when they anticipate to be bailed out by the government later on. As a result, the new bail-in rules might turn out to be not as effective in promoting financial stability as they were originally expected to be. We addressed some of the possible approaches to fix these weaknesses of the "bail-in rules" and concluded that they are unlikely to solve the problem of bank runs on their own due to a combination of asymmetric information and the policy maker's lack commitment.

2.7 Appendix II: Proofs

Proposition 2.4.

Proof. We must show that for every allocation $c_S = (\rho_W, c_{1S}, c_{2S})$ for which both (2.8) and (2.9) is true, we can find a banking contract that would lead to this allocation as the *unique equilibrium* of the withdrawal game for the investors in the bank, conditional on the bank being sound. So suppose that

$$\pi c_{1S} + (1-\pi)\frac{c_{2S}}{R} = 1-\tau$$

$$\rho_S = \pi \quad \text{and} \quad c_{1S} \leq c_{2S}$$

Then given the following contract, the only equilibrium when the bank is sound is for all its investors to follow the no-run strategy.

$$c_1^k\left(0,\,\rho^k\right) = \left\{\begin{array}{c} c_{1S} \\ 0 \end{array}\right\} \quad \text{if} \quad \left\{\begin{array}{c} \rho^k = \pi \\ \rho^k > \pi \end{array}\right\}$$

If all investors follow the no-run strategy (i.e. $x_0^k = 0$) then period-1 requests for withdrawals equal $\rho^k = x_0^k + (1-x_0^k)\pi$ and the contract specifies $c_1^k\left(0,\,\rho^k\right) = c_{1S}$, whereas the budget constraint implies $c_2^k\left(0,\,\rho^k\right) = c_{2S}$, where c_{2S} is obtained from the bank's budget constraint. Since $c_{1S} \leq c_{2S}$ all investors are best responding with the no-run strategy. Specifically, if patient, an investor does not gain from withdrawing in period 1 and therefore best responds by withdrawing in period 2 as specified in the no-run strategy.

On the other hand, for any x_0^k such that $0 < x_0^k \leq 1$, that is, if positive measure of the investors in the bank follow the run strategy in case the bank is sound, then $\rho^k > \pi$ and the above contract would sets $c_1^k\left(0,\,\rho^k\right) = 0$. In this case, the bank's budget constraint in (2.8) yields $c_1^k\left(0,\,\rho^k\right) < c_2^k\left(0,\,\rho^k\right)$. That is, a positive measure $x_0^k(1-\pi) > 0$ of the investors (i.e. those that are patient *and* follow the run strategy) are not best responding. Following an argument analogous to the one above we can establish that a banking contract with the property:

$$c_1^k\left(0,\,\rho^k\right) = \left\{\begin{array}{c} B \\ c_{1S} \\ 0 \end{array}\right\} \quad \text{if} \quad \left\{\begin{array}{c} \rho^k < \rho_S \\ \rho^k = \rho_S \\ \rho^k > \rho_S \end{array}\right\}$$

where $B > c_{2S}$, leads to an allocation satisfying (2.8), and either $\rho_S \in [\pi, 1]$ and $c_{1S} = c_{2S}$ or $\rho_S = 1$ and $c_{1S} \geq c_{2S}$ as the unique equilibrium of the withdrawal game of the bank's investors when the bank's fundamental is sound. □

Proposition 2.5.

Proof. If ρ_S is equal to π then the solution to the problem of maximizing (2.10) subject to (2.8) is given by:

$$u'\left(c_{1S}^{*}\right) = Ru'\left(c_{2S}^{*}\right) \quad \text{and} \quad \pi c_{1S}^{*} + (1-\pi)\frac{c_{2S}^{*}}{R} = 1 - \tau \tag{2.27}$$

Since $R > 1$ we have $c_{1S}^{*} < c_{2S}^{*}$. That is, the allocation $(\rho_S^{*}, c_{1S}^{*}, c_{2S}^{*})$ characterized by (2.27) and $\rho_S^{*} = \pi$ satisfy both (2.8) and (2.9). Next, we show that allocation $(\rho_S^{*}, c_{1S}^{*}, c_{2S}^{*})$, in fact, maximizes the program in (2.10). To see that, consider any other allocation (ρ_S, c_{1S}, c_{2S}) such that $\rho_S = \pi$ and

$$c_{1S} \leq c_{2S} \quad \text{and} \quad \pi c_{1S} + (1-\pi)\frac{c_{2S}}{R} = 1 - \tau$$

That is, the allocation satisfies (2.8) and (2.9) and therefore:

$$V_S\left(c_{1S}^{*},\, \pi\right) > V_S\left(c_{1S},\, \pi\right) \tag{2.28}$$

Next, define the function $\bar{c}(\rho_S)$:

$$\bar{c}(\rho_S) \equiv \frac{1-\tau}{\rho_S + (1-\rho_S)R^{-1}}$$

For each $\rho_S \in [\pi, 1)$ the allocation $(\rho_S,\, \bar{c}(\rho_S),\, \bar{c}(\rho_S))$ satisfies both (2.8) and (2.9) and thus is a potential candidate for the program in (2.10). However, since $\bar{c}(\rho_S)$ is a strictly decreasing function of ρ_S it follows that $V_S\left(\bar{c}(\rho_S),\, \rho_S\right)$ is strictly decreasing in ρ_S. That is,

$$V_S\left(\bar{c}(\pi),\, \pi\right) > V_S\left(\bar{c}(\rho_S),\, \rho_S\right) \tag{2.29}$$

Combining (2.28) and (2.29) yields the desired result, namely the allocation in (2.27) maximizes the function (2.10) subject to (2.8) and (2.9). □

Proposition 2.6.

Proof. Suppose that in equilibrium bank-k's contract is given by c_1^k and the fraction of the bank's investors following the run strategy is $x_{\sigma_k}^k$ for $\sigma_k \in \{0,\, \bar{\sigma}\}$. If the bank is sound, $\sigma_k = 0$, then investors' withdrawal strategies imply that the request for period-1 withdrawals will be equal to

$$\rho_0^k = x_0^k + (1 - x_0^k)\pi \geq \pi \tag{2.30}$$

Given ρ_0^k, the bank's contract would specify a period-1 payment of $c_1^k\left(\rho_0^k,\, 0\right)$ and a period 2 payment of

$$c_2^k\left(\rho_0^k,\, 0\right) = \frac{R\left[1 - \tau - \rho_0^k c_1^k\left(\rho_0^k,\, 0\right)\right]}{1 - \rho_0^k} \tag{2.31}$$

The resulting allocation will be consistent with equilibrium if it also satisfies the implementability constraint in (2.9). That is

$$\rho_0^k = \left\{\begin{array}{c} \pi \\ \in [\pi, 1] \\ 1 \end{array}\right\} \quad \text{as} \quad c_1^k\left(\rho_0^k,\, 0\right) \left\{\begin{array}{c} \leq \\ = \\ \geq \end{array}\right\} c_2^k\left(\rho_0^k,\, 0\right) \tag{2.32}$$

The sum of utilities for the investors in the bank associated with the allocation $\left(\rho_0^k,\, c_1^k\left(\rho_0^k,\, 0\right),\, c_2^k\left(\rho_0^k,\, 0\right)\right)$ will be given by:

$$\rho_0^k u\left(c_1^k\left(\rho_0^k,\, 0\right)\right) + (1 - \rho_0^k) u\left(c_2^k\left(\rho_0^k,\, 0\right)\right) \tag{2.33}$$

Applying Lemma 1, the expression in (2.33) subject to (2.30) - (2.32) is maximized only if the bank's allocation is characterized by $\rho_S^* = \pi$ and (2.27). The allocation in (2.27) will be implemented if the bank's contract is:

$$c_1^k\left(0,\, \rho^k\right) = \left\{ \begin{array}{c} c_{1S}^* \\ 0 \end{array} \right\} \quad \text{for} \quad \left\{ \begin{array}{c} \rho^k = \pi \\ \rho^k \neq \pi \end{array} \right\} \tag{2.34}$$

Indeed, Proposition 2.4 implies that above contract leads to the allocation in Lemma 2 as the *unique equilibrium* when the bank is sound. Hence we obtain the desired result, namely, sound banks do not experience a run $\rho_0^k = \pi$ (i.e. $x_0^k = 0$) and their consumption allocation (c_{1S}^*, c_{2S}^*) is the same as in the constrained efficient case. □

Proposition 2.7.

Proof. Consider an allocation $(\rho_W,\, c_{1W})$ which satisfies the implementability constraint in (2.20). We must consider three cases: (i) $\rho_W = \pi$ and $c_{1W} \leq \hat{c}_2$. (ii) $\pi < \rho_W < 1$ and $c_{1W} = \hat{c}_2$. (iii) $\rho_W = 1$ and $c_{1W} \geq \hat{c}_2$. If $(\rho_W,\, c_{1W})$ is in case (i), that is, $\rho_W = \pi$ and $c_{1W} \leq \hat{c}_2$ then consider the following contract for bank-k

$$c_1^k\left(\bar{\sigma},\, \rho^k\right) = \left\{ \begin{array}{c} c_{1W} \\ 0 \end{array} \right\} \quad \text{for} \quad \left\{ \begin{array}{c} \rho^k = \pi \\ \rho^k > \pi \end{array} \right\} \tag{2.35}$$

If $x_{\bar{\sigma}}^k = 0$, i.e. if all investors in bank-k follow the no-run strategy, then $\rho^k = \pi$ and $c_1^k\left(\bar{\sigma},\, \rho^k\right) = c_{1W}$ and since $c_{1W} \leq \hat{c}_2$ it follows that all investors in bank-k are best responding with the no-run strategy.

On the other hand, for any $x_{\bar{\sigma}}^k > 0$, i.e. if the measure of investors in bank-k following the run strategy is positive, then $\rho^k = x_{\bar{\sigma}}^k + (1 - x_{\bar{\sigma}}^k)\pi > \pi$ and its contract specifies $c_1^k\left(\bar{\sigma},\, \rho^k\right) = 0 < \hat{c}_2$, which violates the the implementability constraint in (2.20). In particular, a fraction $x_{\bar{\sigma}}^k(1 - \pi)$ of the investors in the bank (those that are both patent and follow the run strategy) will not be best responding.

Then, conditional on the bank being weak, an allocation such that $\rho_W = \pi$ and $c_{1W} \leq \hat{c}_2$ will be uniquely implemented by the contract in (2.35).

If (ρ_W, c_{1W}) is in case (ii), that is, $\pi < \rho_W < 1$ and $c_{1W} = \hat{c}_2$. Consider the following contract for bank-k

$$c_1^k \left(\bar{\sigma}, \rho^k\right) = \left\{ \begin{array}{c} \hat{c}_2 + \epsilon \\ c_{1W} \\ 0 \end{array} \right\} \quad \text{for} \quad \left\{ \begin{array}{c} \rho^k < \rho_W \\ \rho^k = \rho_W \\ \rho^k > \rho_W \end{array} \right\} \tag{2.36}$$

If $x_{\bar{\sigma}}^k = z$, where z is such that $\rho^k = z + (1-z)\pi = \rho_W$, we have $c_1^k \left(\bar{\sigma}, \rho^k\right) = c_{1W}$. Since $c_{1W} = \hat{c}_2$ all patient investors are indifferent between withdrawing in period 1 and period 2 and therefore they best respond with the run strategy.

On the other hand, for any $x_{\bar{\sigma}}^k \neq z$ we have $c_1^k \left(\bar{\sigma}, \rho^k\right) = \hat{c}_2 + \epsilon > \hat{c}_2$ for $\rho^k = z + (1-z)\pi < \rho_W$ and $c_1^k \left(\bar{\sigma}, \rho^k\right) = 0 < \hat{c}_2$ for $\rho^k = z + (1-z)\pi > \rho_W$. In either case, $x_{\bar{\sigma}}^k \neq z$ will not be consistent with equilibrium.

Hence, we have shown that conditional on the bank being weak, an allocation such that$\pi < \rho_W < 1$ and $c_{1W} = \hat{c}_2$ will be uniquely implemented by the contract in (2.36).

Finally, if (ρ_W, c_{1W}) is in case (iii), that is, $\rho_W = 1$ and $c_{1W} \geq \hat{c}_2$. Consider the following contract for bank-k

$$c_1^k \left(\bar{\sigma}, \rho^k\right) = \left\{ \begin{array}{c} \hat{c}_2 + \epsilon \\ c_{1W} \end{array} \right\} \quad \text{for} \quad \left\{ \begin{array}{c} \rho^k < 1 \\ \rho^k = 1 \end{array} \right\} \tag{2.37}$$

In this case, we can readily verify that the only equilibrium when outcome k is weak will be for all its investors to follow the run strategy $x_{\bar{\sigma}}^k = 1$, which implies the allocation such that $\rho_W = 1$ and $c_{1W} \geq \hat{c}_2$ is uniquely implemented by the contract in (2.37). $\square$

Proposition2.8.

Proof. Since $\hat{c}_1 < \hat{c}_2$ and $\dfrac{\partial \hat{\rho}_{\bar{\sigma}}}{\partial \rho_W} > 0$ it follows that

$$\frac{\partial V_W\left(c_{1W},\, \rho_W\right)}{\partial c_{1W}} > 0 \quad \text{and} \quad \frac{\partial V_W\left(c_{1W},\, \rho_W\right)}{\partial \rho_W} < 0 \tag{2.38}$$

First, if $c_{1W} < \hat{c}_2$, the implementability constraint (2.20) implies that $\rho_W = \pi$ and from (2.38)

$$V_W\left(c_{1W},\, \pi\right) < V_W\left(\hat{c}_2,\, \pi\right) \quad \text{for} \quad 0 < c_{1W} < \hat{c}_2$$

Second, if $c_{1W} = \hat{c}_2$ then from (2.20) we have $0 \leq \rho_W \leq 1$ and from (2.38)

$$V_W\left(\hat{c}_2,\, \rho_W\right) < V_W\left(\hat{c}_2,\, \pi\right) \quad \text{for} \quad \rho_W > \pi$$

Third, if $c_{1W} > \hat{c}_2$, then (2.20) implies that $\rho_W = 1$ and from (2.38)

$$V_W\left(c_{1W},\, 1\right) < V_W\left(\bar{c},\, 1\right) \quad \text{for} \quad \hat{c}_2 < c_{1W} < \bar{c}$$

where $\bar{c}$ is the maximum payment banks are allowed to make in period 1. The first two inequalities imply

$$V_W\left(c_{1W},\, \rho_W\right) < V_W\left(\hat{c}_2,\, \pi\right) \quad \text{for} \quad c_{1W} \leq \hat{c}_2 \quad \text{and} \quad \rho_W > \pi$$

Therefore, if $\bar{c} < \hat{c}_2$, the bank will set $c_{1W} = \bar{c}$. On the other hand, if $\bar{c} \geq \hat{c}_2$, the bank sets:

(i) $c_{1W} = \bar{c}$ if $V_W\left(\bar{c},\, 1\right) > V_W\left(\hat{c}_2,\, \pi\right)$.

(ii) $c_{1W} \in \{\bar{c},\, \hat{c}_2\}$ if $V_W\left(\bar{c},\, 1\right) = V_W\left(\hat{c}_2,\, \pi\right)$.

(iii) $c_{1W} = \hat{c}_2$ if $V_W\left(\bar{c},\, 1\right) < V_W\left(\hat{c}_2,\, \pi\right)$.

Thus we obtain the desired result, namely c_{1W} will be set as high as possible or equal to $\hat{c}_2$. $\square$

Proposition 2.9.

Proof. The payment in period 1 set by weak banks, c_{1W}, is is either equal to $\bar{c}$ or $\hat{c}_2$. If this were not the case, then Proposition 2.7 implies that banks' contracts are, in fact, not optimal. Also, the maximum payment banks are permitted to set in period-1 is capped at c^*_{1S}, which is the payment in period 1 set by the sound banks (that is, $\bar{c} = c^*_{1S}$). Hence, in equilibrium we have $c_{1W} \in \{c^*_{1S},\, \hat{c}_2\}$. First, suppose that

$$c^*_{1S} > \hat{c}_2$$

and consider the following: (i) weak banks optimize by setting c^*_{1S}. Then since $c^*_{1S} > \hat{c}_2$ their investors best respond with the run strategy ($x_{\bar{\sigma}} = 1 \Rightarrow \rho_W = 1$). According to Proposition 2.7, weak banks would behave optimally if

$$V_W\left(c^*_{1S},\, 1\right) \geq V_W\left(\hat{c}_2,\, \pi\right)$$

An optimal banking contract c^k_1 in this case is given by (2.34) and (2.37), where $c_{1W} = c^*_{1S}$. Another possibility is (ii) weak banks optimize by setting $\hat{c}_2$. Then since $c^*_{1S} > \hat{c}_2$ their investors best respond with the no-run strategy ($x_{\bar{\sigma}} = 0 \Rightarrow \rho_W = \pi$). According to Proposition 2.7, weak banks would behave optimally if

$$V_W\left(c^*_{1S},\, 1\right) \leq V_W\left(\hat{c}_2,\, \pi\right)$$

An optimal banking contract c^k_1 in this case is given by (2.34) and (2.35), with $c_{1W} = \hat{c}_2$. On the other hand, suppose that

$$c^*_{1S} \leq \hat{c}_2$$

then Proposition 2.7 implies that weak banks optimize by setting $c_{1W} = c^*_{1S}$. Also, the fact that $V_W\left(c_{1W},\, \rho_W\right)$ is decreasing in the second argument, yields:

$$V_W\left(c_{1W},\,\pi\right) > V_W\left(c_{1W},\,\rho_W\right)$$

for $\rho_W > \pi$ and therefore an allocation such that $\rho_W > \pi$ is not optimal for banks when $c_{1S}^* \leq \hat{c}_2$. An optimal banking contract c_1^k in this case is given by (2.34) and (2.35), with $c_{1W} = c_{1S}^*$. □

Proposition 2.10.

Proof. The outcome in sound banks is the same both in the run and in the no run equilibrium: these banks do not experience a run and provide the following consumption profile to their investors:

$$u'\left(c_{1S}^*\right) = Ru'\left(c_{2S}^*\right) \quad \text{and} \quad \pi c_{1S}^* + (1-\pi)\frac{c_{2S}^*}{R} = 1-\tau$$

(1) In the no-run equilibrium, weak banks do not experience a run and provide the following consumption profile to their investors:

$$c_{1W} = \min\left\{c_{1S}^*,\,\hat{c}_2\right\} \quad \text{and} \quad \pi\hat{c}_1 + (1-\pi)\frac{\hat{c}_2}{R} = 1-\tau-\bar{\sigma}-\theta c_{1W}+\hat{b}$$

$$u'\left(\hat{c}_1\right) = Ru'\left(\hat{c}_2\right) = v\left(\tau - n\hat{b}\right)$$

And the tax rate is determined by (2.4). The *no-run equilibrium* exists if

$$c_{1W} \leq \hat{c}_2$$

$$\theta u\left(c_{1W}\right) + (\pi-\theta)u\left(\hat{c}_1\right) + (1-\pi)u\left(\hat{c}_2\right) \geq \theta u\left(c_{1S}^*\right) + (1-\theta)\left[\pi u\left(\hat{c}_1\right) + (1-\pi)u\left(\hat{c}_2\right)\right]$$

The first condition ensures that investors in weak banks best respond with the no-run strategy when these banks set c_{1W}, the second condition ensures that weak banks want to set c_{1W}.

(2) In the *run equilibrium* weak banks experience a run and provide the following consumption profile:

$$c_{1W} = c_{1S}^* \quad \text{and} \quad \pi \hat{c}_1 + (1-\pi)\frac{\hat{c}_2}{R} = 1 - \tau - \bar{\sigma} - \theta c_{1S}^* + \hat{b}$$

$$u'(\hat{c}_1) = Ru'(\hat{c}_2) = v\left(\tau - n\hat{b}\right)$$

And the tax rate is determined by the first order condition in (2.4). The *run equilibrium* exist if

$$c_{1S}^* \geq \hat{c}_2$$

$$\theta u(c_{1S}^*) + (1-\theta)\left[\pi u(\hat{c}_1) + (1-\pi)u(\hat{c}_2)\right] \geq \theta u(\hat{c}_2) + (\pi-\theta)u(\hat{c}_1) + (1-\pi)u(\hat{c}_2)$$

The first condition ensures that investors in weak banks best respond with the run strategy when these banks set c_{1S}^*, the second condition ensures that weak banks want to set c_{1S}^*.

For given parameter values (R, π, γ, δ, q, θ, n, $\bar{\sigma}$) and investor preferences as in (2.26) we can explicitly solve for the the allocation in the run and no-run equilibrium respectively and then verify if the resulting allocations are consistent with equilibrium. The result of this calculation is show in Figure 2 which also establishes the statement in the Proposition. □

Proposition 2.11.

Proof. A necessary condition for the equilibrium to be constrained efficient is for weak banks to pay the same amount to all investors withdrawing in period 1 that is $c_{1W} = \hat{c}_1$. That is, from the start of period 1, weak banks must lower their payments all the way down to their level in resolution $\hat{c}_1$. From Proposition 2.5, however, the equilibrium value of c_{1W} would equal either c_{1S}^* or $\hat{c}_2$. Then since $\min\{c_{1S}^*, \hat{c}_2\} > \hat{c}_1$ it follows that the equilibrium is not constrained efficient. $\square$

Chapter 3

Private Sunspots in a Coordinated Attack Game

3.1 Introduction

A number of important economic phenomena such as currency attacks, bank runs and sovereign defaults can be understood and modeled as collective action games where the players can coordinate on one of two outcomes with very different consequences in terms of welfare and policy. Such coordination occurs because of strategic complementarities where the benefit of an action for a given player - attacking the currency, running on the bank, lending to the government – increases with the number of other players choosing the same action. Multiple equilibria play a central role in models of bank runs (Diamond and Dybvig, 1983), currency crises (Obstfeld, 1996) and sovereign defaults (Calvo, 1988). This multiplicity of possible equilibrium outcomes, however, presents a theoretical challenge since it renders the model predictions and its comparative statics relatively ambiguous. The literature has proposed different methods for dealing with this problem. This methods fall into two general categories – the sunspot-based approach and the global game approach.[1] [2]

In this chapter, I propose a novel technique to endogenize the probability of a coordination event within the sunspot-based approach. Specifically, I investigate

[1]The global game framework was initially proposed by Carlson and Van Damme (1993) and further developed by Frankel, Morris, and Pauzner (2002). Applications include currency crises (Morris and Shin, 1998), bank runs (Goldstein and Pauzner, 2005), debt crisis (Morris and Shin, 2004), business cycles (Burdzy and Frankel, 2005), investment cycles (Chamley, 1999, Oyama, 2004), merger waves (Toxvaerd, 2008) and competing computer platforms (Argenziano, 2008). A thorough exposition of the global game approach is provided in Morris and Shin (2002).

[2]The sunspot approach originated with Cass and Shell (1983) and has been used to study a wide range of issues in macroeconomics (Azariadis, 1981; Woodford, 1986), monetary economics (Smith, 1988), learning (Woodford, 1990), business cycles (Benhabib and Farmer, 1994), and bank runs (Cooper and Ross, 1988; Peck and Shell, 2003), among many other topics. For an overview of this literature, see Shell (2008).

what happens when agents receive a noisy signal of a sunspot state. I deliver the central result of the private sunspot approach in the context of a simple example called the Bandit Game. Specifically, there are three players: bandit 1, bandit 2 and the village. The bandits are isolated from each other and want to coordinate on attacking the village with the goal of seizing its output. The village, in turn, must decide how much output to produce. The setup of the Bandit Game shares some of the properties of a classical coordinated attack problem with some important differences. First, the village makes a strategic choice - namely how much to produce - which directly affects the benefit of the attack. Second, the amount produced by the village remains hidden from the bandits up to the point they enter the village (that is, if there is an attack). In equilibrium, of course, each bandit will infer the amount produced and the village, in turn, will correctly anticipate the probability of a coordinated attack. Third, the village cannot commit to a specific level of output ex-ante, but instead best responds to the perceived probability of an attack.

It is well known that in games of coordination the strategies of the players might be conditioned on the realizations of a sunspot state. This sunspot state is not related to the fundamentals or other payoff-relevant factors and represents purely extrinsic uncertainty. Specifically, each bandit can base his choice of whether or not to attack the village on the realizations of the sunspot state if he expects the other to do the same. The probability of a coordinated attack on the village is therefore equal to the probability that this sunspot state takes on values for which the bandits choose to attack. With the usual sunspot-based approach, however, the equilibrium probability of an attack in the Bandit Game is indeterminate: it can be any number between zero and an upper bound. Such a prediction is not satisfactory and appears to reveal a weakness in the sunspot-based approach to coordination games.

The purpose of my investigation is to propose a way to deal with this issue. Specifically, notice that if the sunspot state were perfectly observed, there will be no strategic uncertainty. In other words, given his signal, each bandit can perfectly predict the action that will be undertaken by the other bandit. The idea of the private sunspot approach is to introduce a small degree of strategic uncertainty via the sunspot state. Specifically, I perturb the original game by assuming that each bandit receives his own signal of the realization of the sunspot state (this signal is what is called his *private sunspot*), which is arbitrarily close to the true realization of the sunspot state. In other words, the sunspot state is no longer common knowledge, but almost common knowledge.[3] The resulting perturbed game, thought arbitrarily close to the one without strategic uncertainty, will be shown to have markedly different properties. Specifically, I will show that there will be only one value for the probability of an attack which is both positive and consistent with equilibrium. This value will be pinned down by the parameters of the model.

In general, the introduction of private sunspots is a necessary but not sufficient condition for strategic uncertainty to arise. Specifically, if the bandits choose the same action for each value of their private sunspot (either to attack or not to attack) then there will be no strategic uncertainty. What is true, however, is that as long as there is strategic uncertainty – as will be the case if the bandits take different actions for different values of their private sunspots – then the equilibrium must satisfy an additional condition. This condition will be what allows us to pin down a unique probability of an attack in the presence of strategic uncertainty.[4]

One interpretation of the private sunspot approach is that it allows the bandits to hold idiosyncratic sentiments about the prospect for successful attack. These

[3]Rubinstein (1988) is an early example of the coordinated attack problem in a game of almost common knowledge. His setup, however, is different from the one here.

[4]This result is reminiscent to the global game literature where the equilibrium of the original game of coordination is not robust to a small perturbation that introduces noisy signals with respect to fundamentals (see Kajii and Morris (1997) for a theory of "robustness to incomplete information"). In a global game application, this noisy approximation is adding both fundamental and strategic uncertainty to the original complete information formulation of the game. On the other hand, in the private sunspot case, the level of fundamental uncertainty would remain unchanged and the small perturbation adds only strategic uncertainty.

idiosyncratic sentiments are generated by the imperfect observations of a variable which is not related to the fundamentals of the environment, but which nonetheless conveys information about the likely action of the other bandit. In other words, their degree of optimism would be determined endogenously, and would not necessarily coincide, even though their information about the fundamentals and any other payoff-relevant features of the environment is always the same. For example, one might imagine that the bandits' coordination device (i.e. the sunspot) is noisy and not entirely reliable because its realizations cannot be measured exactly and/or is open to interpretation.

This last point has been made by Angeletos (2008) who analyses a model with private sunspots. There are, however, important differences between his approach and the analysis presented here. First, there is no agent playing the role of the village in Angeletos (2008). In contrast, it will soon become obvious that the village's strategic choice is central to my analysis. Specifically, the equilibrium actions of the village will allow us to pin down the probability of an attack. Second, Angeletos is primarily interested in the way private sunspots induce heterogeneous investor sentiment and in the variation in the equilibrium actions even if all players share the same information with respect to fundamentals and other payoff-relevant outcomes.[5] The analysis here, in contrast, aims to approximate the original game by introducing small noise in the player's coordination device. Ex-ante, the probability that the bandits will choose different actions will, in fact, be arbitrarily close to zero.

The rest of this chapter is organized as follows. Section 3.2 introduces the Bandit Game and characterizes its Nash equilibria. The main result of this chapter is established in Section 3.3 which introduces two types of sunspot equilibria – one where the sunspot state is perfectly observed and one where it is not. The private sunspot approach is based on logic similar to that of the global games framework and therefore it

[5] See also Gu (2011), which investigates how observing noisy, private sunspot signals influences the withdrawal decisions of depositors in a model of bank runs.

will be instructive to examine a Global Games version of the bandits' game with the purpose of further illuminating the properties of the private sunspot approach. This is done in section 3.4, where I compare the private sunspots approach to the Bandit Game with the global game approach. In order to do that, I first recast the Bandits Game as a global game, then derive the equilibrium probability of a successful attack and finally compare its properties with that of the private sunspots. In section 3.5, I show that the logic of the Bandit Game can be readily generalized for a general class of coordination games involving a set of players (i.e. the bandits, speculators, depositors) making a binary choice (i.e. whether or not to attack, short sell the currency, run on the bank) and where the benefit derived from successful coordination is a function of the choice made by another player (i.e. the village, the government, the bank). Section 3.6 concludes and outlines directions for future research.

3.2 The Bandit Game

There are three players: bandit 1, bandit 2 and the village. The village chooses how much output to produce. In order to produce an output of θ, the village must pay a cost of $t\theta^2/2$. The bandits are isolated from each other and do not observe the output produced by the village (i.e. it is hidden in the village). Each bandit must decide whether to attack or not to attack the village. For given θ, the payoff matrix for the bandits will be the following:

	Attack	Not attack
Attack	$\theta/2 - c$, $\theta/2 - c$	$-c$, 0
Not attack	0, $-c$	0, 0

Table 3.1: Payoff Matrix for given θ.

A bandit that attacks the village must pay a cost of $c > 0$. If both bandits attack, they split whatever output the village had produced equally between them. If only one of then attacks, he is unsuccessful and pays the cost of the attack c. For example, one can imagine that the output θ is well guarded and impossible for a bandit to seize on his own. If the village chooses to produce θ, then it consumes θ unless both bandits attack, in which case the village consumes 0. Thus, if fewer than two bandits attack, the payoff to the village is $\theta - t\theta^2/2$, whereas, if both bandits attack, the payoff to the village is $-t\theta^2/2$.

3.2.1 Nash equilibria of the Bandit Game

I begin by deriving the set of Nash Equilibria of the Bandit Game. A strategy for each bandit is $p_i \in [0,1]$, where p_i is the probability that bandit i attacks the village. The probability of a coordinated attack on the village is the probability that both bandits choose to attack, namely p_1p_2. A strategy for the village is a choice of output $\theta \geq 0$. The village chooses its output level in order to maximize its expected payoff

$$\underset{\theta}{max} \quad (1 - p_1p_2)\theta - t\frac{\theta^2}{2}$$

The level of output produced by the village depends on the anticipated probability of a coordinated attack and is equal to $\theta = (1 - p_1p_2)/t$.

Proposition 3.1. ***Nash equilibria of the Bandits Game***

If $ct > \frac{2}{27}$, then the Bandits Game has a unique equilibrium where both bandits do not attack and the village's output is $1/t$.

If $ct < \frac{2}{27}$, then the bandits game has three equilibria; one in pure strategies where both bandits do not attack and the village's output is $1/t$ and two mixed strategy equilibria.

Proof. If bandit 1 is not attacking (i.e. $p_1 = 0$), then the best response for bandit 2 will be not to attack as well (and vice versa). Thus there exist an equilibrium where both bandits do not attack $p_1 = p_2 = 0$ and the village sets $\theta = 1/t$. On the other hand, an equilibrium where both bandits attack with probability 1 does not exist.

Indeed, if $p_i = 1$ for $i = 1, 2$, the village would best respond by $\theta = 0$. But if $\theta = 0$, each bandit will best respond by choosing not to attack. Hence there does not exist an equilibrium with $p_1 = p_2 = 1$. Next, suppose that both bandits are mixing by choosing to attack with probability $p_i \in (0, 1)$. Each bandit is best responding if and only if:

$$p_i \left(\frac{\theta}{2} - c \right) + (1 - p_i)(-c) = 0$$

which implies that the the mixing probabilities must satisfy $p_i = \dfrac{2c}{\theta}$ for $i = 1, 2$. By combining with the village's first order condition for the choice of θ, it follows that an equilibrium mixing probability is determined as the real roots of the polynomial $f(p)$ in the interval $[0, 1]$, where

$$f(p) \equiv p^3 - p^2 + 2ct$$

If $ct < 2/27$ then f has two real roots in $[0, 1]$, whereas if $ct > 2/27$ then f has no real roots in $[0, 1]$.[6] $\square$

Figure 3.1 plots $f(p)$ for $c = t = 1/5$. There is an equilibrium in pure strategies where neither bandit attacks, $p = 0$. In addition, there are two equilibria in mixed strategies, at $p = 0.35$ and $p = 0.9$.

[6]Since $f(0) = f(1) = 2ct$ the function f has two roots in the interval $(0, 1)$ whenever $f(p_L) < 0$ where, $f(p_L)$ is the lowest value of the function f in the interval $[0, 1]$ which is attained at $p_L = \dfrac{2}{3}$. Thus, $f\left(\dfrac{2}{3}\right) < 0 \Leftrightarrow ct < \dfrac{2}{27}$.

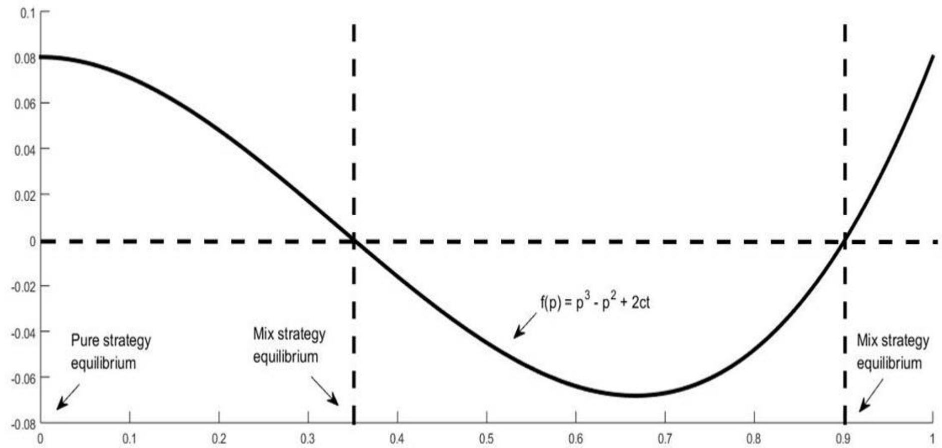


Figure 3.1: Nash equilibria of the Bandit Game

3.3 Sunspot equilibria of The Bandits Game

If we expand out concept of equilibrium beyond Nash, other equilibrium outcomes can arise as well. For example, the strategies of the bandits might be conditioned on the realization of a sunspot state. This sunspot state is external to the fundamental of the model and serves as a coordination device. Specifically, if there is a set of mutually agreed values of the sunspot for which both bandits attack, then the probability of an attack on the village will be equal to the probability that the sunspot takes on one of these values. Such equilibria will be examined next. Specifically, I will examine two types of sunspot equilibria: one where the sunspot state is observed perfectly by the bandits and one where the sunspot state is not observed and each bandit receives a noisy private signal of its realization (his *private sunspot*).

3.3.1 The sunspot state

The sunspot state s is the realization of a payoff-irrelevant random variable which has a uniform distribution on the unit interval $s \sim U[0,1]$. The distribution of this sunspot state is assumed to be common knowledge for all players in the game. After the realization of the sunspot state s each bandit receives a private signal:

$$s_i = s + \epsilon_i \tag{3.1}$$

where ϵ_i is uniformly distributed noise term $\epsilon_i \sim U[-\epsilon, \epsilon]$. Hence, the signal for each bandit (i.e. his private sunspot) is distributed uniformly around the realization of the sunspot state:

$$s_i \,|s \sim U[s - \epsilon, s + \epsilon] \tag{3.2}$$

Conditional on s_i, bandit i posterior beliefs about the sunspot state and about the signal received by the other bandit will be the following:

$$s\,|s_i \sim U[s_i - \epsilon, s_i + \epsilon] \quad \text{and} \quad s_{-i}\,|s_i \sim U[s_i - 2\epsilon, s_i + 2\epsilon] \tag{3.3}$$

The village does not receive a signal and hence continues to hold the prior for the sunspot state, that is $s \sim U[0, 1]$.

The bandits can now follow strategies which are contingent on their signals. In particular, a strategy for bandit i is a mapping from signals to actions: $a_i : S \to \{0, 1\}$, where $a_i = 1$ corresponds to attacking the village and $a_i = 0$ corresponds to not attacking the village.

Definition 3.2. An *attack set* $A \subseteq [0, 1]$ is such that $a_i(s_i) = 1$ if and only if $s_i \in A$ for $i = 1, 2$.

I restrict attention to attack sets characterized by $n \geq 0$ threshold points. If the attack set has no threshold points $n = 0$, then the bandits take the same action for each value of their private signal s_i. In particular, if $A = [0, 1]$ they choose to attack for each realization of their private signal. That is, $a_i(s_i) = 1$ for all s_i.In contrast, if $A = \emptyset$ then they choose not to attack for each realization of their private signal, i.e.,

$a_i(s_i) = 0$ for all s_i. On the other hand, suppose that the attack set A is characterized by a finite number n of threshold points $\{s_k^*\}_{k=1}^n$ such that:

$$0 < s_1^* < s_2^* < ... < s_n^* < 1$$

Each threshold point signifies a change from one action to another. Thus, if the attack set A is characterized by a single threshold point s^*, then:

$$A = [s^*,\, 1] \quad \text{or} \quad A = [0,\, s^*]$$

In the first case $a_i(s_i) = 1$ iff $s_i \geq s^*$, that is, each bandit attacks only if his signal is greater than or equal to the threshold point s^*. In the second case $a_i(s_i) = 1$ iff $s_i \leq s^*$, i.e., each bandit attacks only if his signal is less than or equal to the threshold point s^*. On the other hand, if the attack set A is characterized by two threshold points s_1^* and s_2^*, then

$$A = [0,\, s_1^*] \cup [s_2^*,\, 1] \quad \text{or} \quad A = [s_1^*,\, s_2^*]$$

and so on for any given n. Given their strategies, the ex-ante probability that both bandits coordinate and choose to attack the village is equal to the probability that they both receive a signal in A. In the limit as the noise in the private sunspot converges to zero, the probability of a joint attack on the village converges to the probability that the sunspot state belongs to A. That is,

$$\underset{\epsilon \to 0}{lim} \Pr[s_1 \in A,\ s_2 \in A] = \Pr[s \in A] \tag{3.4}$$

For example, if the attack set is characterized by a single threshold point such that (for example $A = [0,\, s^*]$) then the probability of an attack on the village is equal to the probability the sunspot state is less than s^* i.e. $\Pr[s \leq s^*] = s^*$.

I will examine two types of sunspot equilibria, depending on whether the realization of the sunspot state is observed perfectly by the bandits ($\epsilon = 0$) or with a small noise ($\epsilon > 0$).

3.3.2 The sunspot state is observed perfectly

In this section, I assume that the sunspot state is observed perfectly by each bandit, meaning $\epsilon = 0$. In this case there will be no strategic uncertainty. In other words, given his signal, each bandit can perfectly predict the action that will be undertaken by the other bandit in equilibrium.

Proposition 3.3. *If the sunspot state is observed perfectly by the bandits, then for each* $0 \leq q \leq 1 - 2tc$ *there exists a sunspot equilibrium where the probability of a coordinated attack is* q *and the output produced by the village is* $\theta = \dfrac{1-q}{t}$.

Proof. The probability of an attack will be equal to $q \in [0, 1]$ if the bandits follow a threshold strategy of the form

$$a_i(s) = 1 \quad \text{iff} \quad s \leq s^*(q) \equiv q$$

Fix q and suppose that the bandits attack the village if and only of the sunspot state s is less or equal to $s^*(q)$. The probability of an attack on the village is thus equal to q and the best response choice of output for the village is $\theta(q) = (1-q)/t$. Each bandit, in turn, will be best responding by attacking whenever $\theta(q)/2 - c \geq 0$, which is equivalent to $0 \leq q \leq 1 - 2tc$. □

According to Proposition 3.3 we can find an equilibrium such that the probability of a coordinated attack on the village is equal to any value in the interval $0 \leq q \leq 1 - 2ct$. The set of equilibria in this case are shown on Figure 3.2. The probability of an attack on the village is on the horizontal axis and the best response output level is on the vertical axis. Observe that not all values of q can be supported in equilibrium since for q sufficiently close to 1 the village's output θ will be not enough to cover the cost of an attack c.

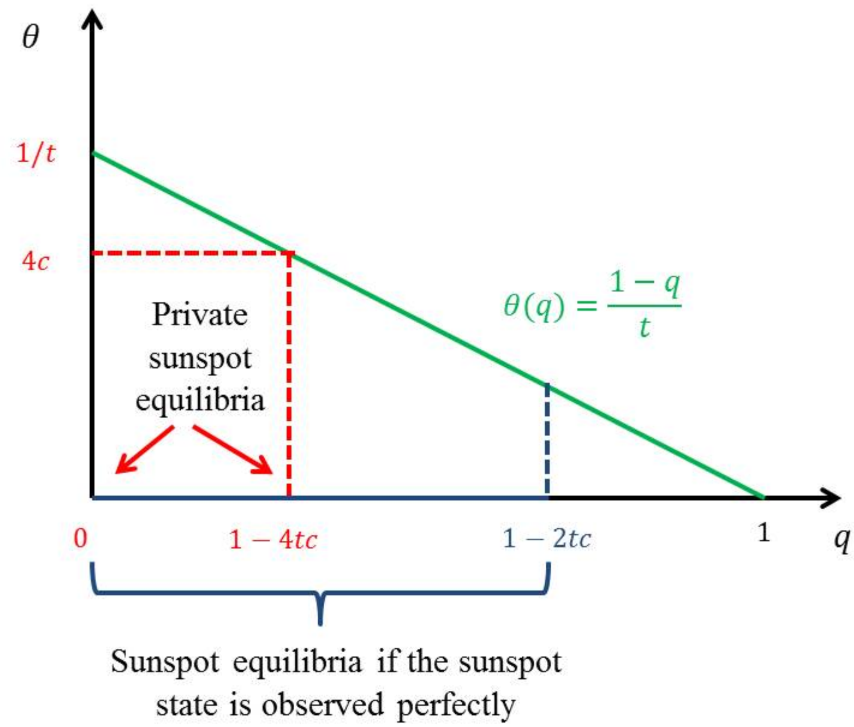


Figure 3.2: Sunspot equilibria of the Bandits Game

3.3.3 Private sunspots

Observe that the two bandits always make the same equilibrium choice if the sunspot state is perfectly observed. Their actions, in contrast, would not necessarily be the same in the private sunspot case where their signals of the sunspot state have a small noise $\epsilon > 0$. Indeed, if the bandits follow a threshold strategy and bandit i's signal is in close proximity to the threshold point s^*, then he will be uncertain as to the action of the other bandit. This strategic uncertainty will be shown to play a vital role in the private sunspot approach by imposing an additional restriction that must be satisfied in equilibrium. The main result of the paper is established in Proposition 3.4. Specifically, I show that the introduction of small noise in the signals for the bandits would reduce the set of q consistent with equilibrium to only two values - one equal to zero and the other positive.

Proposition 3.4. *If the bandits observe the sunspot state with a small noise, then there are only two values for the probability of a coordinated attack which are consistent with equilibrium:* $q^{\mathrm{NA}} = 0$ *and* $q^{\mathrm{P}} = 1 - 4tc$.

Proof. Consider a profile of strategies for the bandits characterized by an attack set $A \subseteq [0, 1]$:

$$a_i(s_i) = 1 \quad \text{iff} \quad s_i \in A \quad \text{for} \quad i = 1, 2 \tag{3.5}$$

This strategy profile generates a probability of an attack of $q = \Pr[s_1 \in A,\, s_2 \in A]$, which in the limit as the noise in their private signals goes to zero converges to $q = \Pr[s \in A]$.

First, an attack with probability $1 \geq q > 1 - 2ct$ cannot be sustained in equilibrium, since in this case, the output selected by the village $\hat{\theta}(q) = \dfrac{1-q}{t}$ would imply $\dfrac{\hat{\theta}(q)}{2} - c < 0$. But this makes choosing not to attack a best response for each bandit, regardless of his signal. Hence, in equilibrium we must have $q \leq 1 - 2ct$.

Second, suppose that both bandits choose not to attack for each value of their private signals, that is $q = 0$. Then each bandit is best responding with $a_i = 0$ (since choosing to attack when the other is expected not to is never optimal) and therefore $q = 0$ (with associated output of $\hat{\theta}(0) = 1/t$) is consistent with equilibrium.

Third, consider values of q in the interval $0 < q \leq 1 - 2ct$. I show that the only q consistent with equilibrium in this interval equals $1 - 4tc$. If $0 < q < 1$, the attack set A must contain at least one threshold point. The remaining of the proof is organized in two steps. First, I assume that A has a single threshold point and show that the only value of q consistent with equilibrium in the interval $(0,\, 1-2tc]$ must equal $1-4tc$. Second, I allow for A to contain an arbitrary number of threshold points and still show that the only positive q that can be sustained in equilibrium equals $1-4tc$.

Step 1. Suppose that the attack set A has a single threshold point $n = 1$. For concreteness, consider a profile of strategies of the form:

$$a_i(s_i) = 1 \quad \text{iff} \quad s_i \geq s^* \quad \text{for} \quad i = 1, 2 \tag{3.6}$$

That is, each each bandit follows a threshold strategy and attacks whenever his signal greater or equal to s^*.[7] The probability of a coordinated attack on the village thus equals the probability that each bandit receives a signal above s^*. In the limits as the noise in their signals goes to zero this probability becomes simply $q = \Pr[s \geq s^*]$. Given the signal for bandit i, the posterior probability that the other bandit would attack is equal to the probability of the event $s_{-i} \geq s^* \,|s_i$ which is obtained from:

$$\Pr[s_{-i} \geq s^* \,|s_i] = \begin{cases} 0 & s_i < s^* - 2\epsilon \\ \dfrac{1}{2} + \dfrac{s_i - s^*}{4\epsilon} & \text{if} \quad s^* - 2\epsilon \leq s_i \leq s^* + 2\epsilon \\ 1 & s_i > s^* + 2\epsilon \end{cases} \tag{3.7}$$

The above function is illustrated on the left panel in Figure 3.3. For given s_i, the expected net gain for bandit i of choosing $a_i = 1$ (to attack) relative to choosing $a_i = 0$ (not to attack) is the following:

$$\Delta(s_i) \equiv \Pr[s_{-i} \geq s^* \,|s_i] \left(\frac{\hat{\theta}(q)}{2} - c \right) + (1 - \Pr[s_{-i} \geq s^* \,|s_i])\,(-c) \tag{3.8}$$

Each bandit will be best responding with the strategy in (3.6) if he prefer to attack if $s_i > s^*$, not to attack if $s_i < s^*$ and is indifferent between the two actions if $s_i = s^*$. That is,

[7] Here I will restrict attention to $A = [s^*,\, 1]$. The case $A = [0,\, s^*]$ represents a mirror image and can be treated analogously.

$$\Delta(s_i)\left\{\begin{array}{c} > \\ = \\ < \end{array}\right\}0 \quad \text{if} \quad s_i\left\{\begin{array}{c} < \\ = \\ > \end{array}\right\}s^* \tag{3.9}$$

The function $\Delta(s_i)$ is shown on the right panel in Figure 3.3. From (3.7) we obtain that

$$\Pr\left[s_{-i} \geq s^* \left| s^* \right.\right] = \frac{1}{2}$$

That is, a bandit whose signal equals the threshold assigns probability 1/2 to the event that the other bandit received a signal above s^* and will choose to attack. The above condition, combined with (3.9), also implies that a bandit whose signal is equal to the threshold point must be indifferent between his two actions when the other bandit is mixing with equal probability. That is,

$$\Delta\left(s^*\right) = \frac{1}{2}\left(\frac{\hat{\theta}\left(q^{\mathrm{P}}\right)}{2} - c\right) + \frac{1}{2}(-c) = 0 \tag{3.10}$$

Condition (3.10) holds if and only if the output chosen by the village is such that:

$$\hat{\theta}\left(q^{\mathrm{P}}\right) = \frac{1-q^{\mathrm{P}}}{t} = 4c \tag{3.11}$$

Thus, we obtain the probability of a coordinated attack must be equal to $q^{\mathrm{P}} = 1-4tc$. Any other value in the interval $(0,\, 1-2tc]$ implies $\hat{\theta}(q) \neq 4c$ and therefore would violate the condition in (3.10).

Step 2. Suppose that the attack contains $n \geq 2$ threshold points. Around each threshold point s_k^*, the bandits will be switching their action from an attack $a_i = 1$ to not attack $a_i = 0$ or vice versa.

First, let x denote an arbitrary threshold such that the bandits are switching from not attacking $a_i = 0$ to attacking $a_i = 1$ (such as point s_{k+1}^* on Figure 3.4 and point

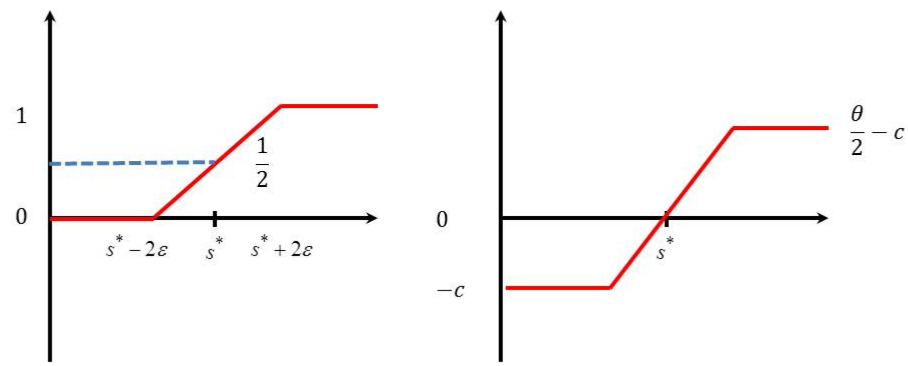


Figure 3.3: The bandit strategies are characterized by only one threshold point s^* such that each bandit attacks if and only if $s_i \geq s^*$.

s^*_k on Figure 3.5). Since $n \geq 2$, the attack set must contain at least one threshold point with this property. In a small neighborhood of such a point x, the following must be true in equilibrium:

$$\Pr\left[A\,|s_i \in O(x)\right] = \begin{cases} 0 & s_i < x - 2\epsilon \\ \dfrac{1}{2} + \dfrac{(x - s_i)}{4\epsilon} & \text{if} \quad x - 2\epsilon \leq s_i \leq x + 2\epsilon \\ 1 & s_i > x + 2\epsilon \end{cases} \tag{3.12}$$

$$\mathit{\Delta}(s_i) \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 0 \quad \text{if} \quad s_i \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} x \quad \text{and} \quad s_i \in O(x) \tag{3.13}$$

where $O(x) \equiv \{s_i \quad \text{s.t} \quad s_i \in [x - r,\, x + r]\}$ and $r > 0$ is small enough to ensure that $O(s^*_k)$ and $O(s^*_{k+1})$ do not overlap for all k. This is illustrated on Figures 3.4 (with respect to s^*_{k+1}) and on Figure 3.5 (with respect to s^*_k).

Second, let y be a threshold point such that the bandits switch their action from attacking $a_i = 1$ to not attacking $a_i = 0$ (such as point s^*_k on Figure 3.4 or point s^*_{k+1} on Figure 3.5). Since $n \geq 2$, the attack set must contain at least one threshold point with this property. The following must be true in equilibrium:

$$\Pr\left[A \,|s_i \in O(y)\right] = \begin{cases} 1 & s_i < y - 2\epsilon \\ \dfrac{1}{2} - \dfrac{(y - s_i)}{4\epsilon} & \text{if} \quad y - 2\epsilon \leq s_i \leq y + 2\epsilon \\ 0 & s_i > y + 2\epsilon \end{cases} \tag{3.14}$$

$$\Delta(s_i) \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} 0 \quad \text{if} \quad s_i \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} y \quad \text{and} \quad s_i \in O(y) \tag{3.15}$$

where $O(y) \equiv \{s_i \quad \text{s.t} \quad s_i \in [y - r,\, y + r]\}$ is a small neighborhood around the point y. From (3.12) - (3.13) and (3.14) - (3.15) we obtain that the bandit whose signal equals any of the threshold points must be indifferent between choosing to attack and choosing not to attack if the other bandit is mixing with equal probability.

$$\Delta(s_i) = \frac{1}{2}\left(\frac{\hat{\theta}(q)}{2} - c\right) + \frac{1}{2}(-c) = 0 \quad \text{if} \quad s_i \in \{s^*_k\}_{k=1}^{n} \tag{3.16}$$

If $q > 0$ then in equilibrium the bandits are best responding if and only if the condition in (3.16) is satisfied. Notice that this condition is independent of the number of threshold points (or their specific values). Moreover, this is precisely the same condition as in (3.10), where the bandits strategies had only one threshold point. Thus, we must have $\hat{\theta}(q) = 4c$, which will be the case if and only if $q^{\mathrm{P}} = 1 - 4tc$. In this sense, a set with a single threshold point is just a normalization and thus does not entail any loss of generality. $\square$

If the sunspot state were perfectly observed, then any value of q in the interval $[0, 1-2ct]$ can be sustained in equilibrium (Proposition 3.3). In contrast, if the sunspot state is observed almost perfectly, then there is only one value of q which is both positive and consistent with equilibrium, namely $q^{\mathrm{P}} = 1-4tc$ (Proposition 3.4). Indeed, if the probability of an attack were greater (less) than $q^{\mathrm{P}} = 1-4tc$, then the village's choice of output $\hat{\theta}(q)$ will be less (greater) than $4c$. But if that were the case, then the condition in (3.16) will be violated. Thus, we are left with two options for the equilibrium probability of a coordinated attack on the village. Either $q = 0$, in which case the bandits' strategies have no threshold points and the condition in (3.16) need not be satisfied, or the equilibrium probability if an attack is positive and given by $q^{\mathrm{P}} = 1-4tc$. Hence, the condition in (3.16) provides an additional restriction that allows us to endogenize the probability of an attack by linking it directly to the parameters of the model.[8]

[8] Also, note that this condition in (3.16) does not depend directly on the ex-ante probability of an attack on the village q since the bandits' base their choices on their more precise private signals.

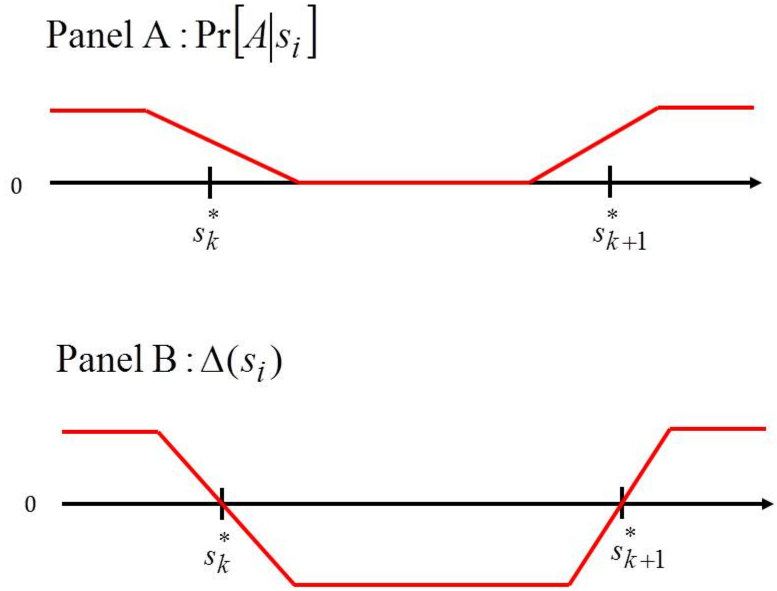


Figure 3.4: The bandit strategies are characterized by at least two threshold points.

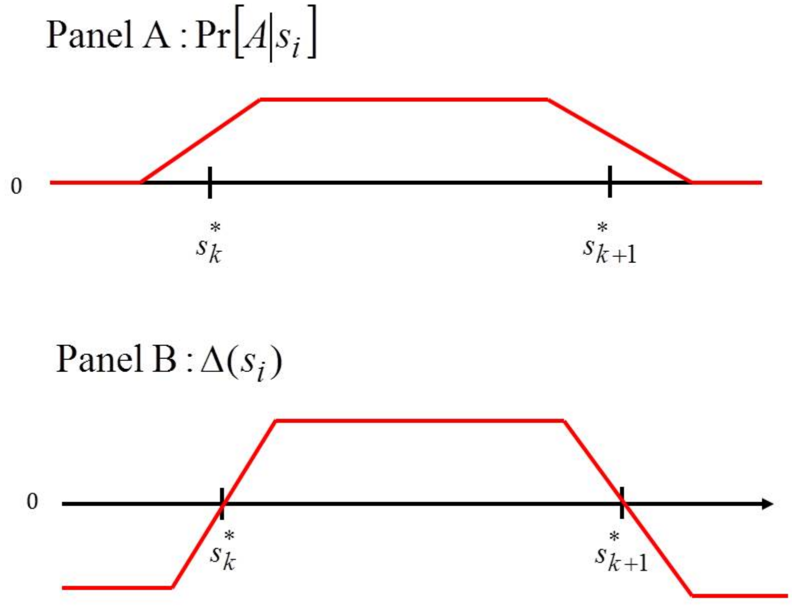


Figure 3.5: The bandit strategies are characterized by at least two threshold points.

3.3.4 Properties of the private sunspots equilibrium

The private sunspot equilibria - though based entirely on self-fulfilling expectations - nevertheless leads to markedly different predictions with respect to the probability of an attack on the village compared to the case where the sunspot is observed perfectly. In this section, I further examine some of the properties of the private sunspot equilibria.

Strategic uncertainty. The fact that the sunspot is imperfectly observed is only a necessary, but not a sufficient condition for strategic uncertainty. For example, there will be no strategic uncertainty if the bandits always choose the same action for each value of their private sunspot. In contrast, if the bandits switch their action around a threshold point, then there is strategic uncertainty. Specifically, if a given bandit receives a signal in a close proximity to this threshold point, he will be unable to perfectly predict the action the other bandit will take (given that the other bandit signal might either be above or below the threshold). This strategic uncertainty would impose an additional equilibrium condition that must be satisfied by the bandits' strategies, namely the one in (3.16). Specifically, a bandit whose signal equals the threshold must be indifferent between his two actions while assigning an equal probability that the other bandit received a signal above and below this threshold. This condition would not appear if the sunspot state were perfectly observed, for in this case, each bandit knows precisely the equilibrium action of the other bandit.

Relation to risk dominance. If sunspots are private and there is strategic uncertainty (as will be the case if bandits do not always take the same action) then in equilibrium the condition in (3.16) must hold. This, however, will be the case **only if** the level of output chosen by the village equals $4c$. The village's choice, at the same time, depends on the ex-ante probability of an attack.[9] Thus what is necessary in equilibrium is for the probability of an attack q^{P} to be such that the village's choice is $\hat{\theta}\left(q^{\mathrm{P}}\right) = 4c$. The value of output ensuring that the condition in (3.16) is satisfied is important and deserves further emphasis. One can relate the results in this section to the notion of risk dominance in the sense of Harsanyi and Selton (1998). Specifically, suppose we fix the level of output to θ and consider the complete information subgame associated with this value of θ. In this case, attacking will be the risk dominant action if $\theta > 4c$, not attacking will be the risk dominant action if $\theta < 4c$.

The sunspot state. The assumption that both the sunspot state and the signals are uniform, that is, $s \sim U\,[0, 1]$ and $s_i\,|s \sim U\,[s - \epsilon,\, s + \epsilon]$ is without loss of generality and was made for simplicity. In general, one can consider a sunspot state s with a domain S and a cumulative distribution function F. Given the realization of s, each bandit receives an unbiased and bounded signal s_i with domain in $[s - \epsilon,\, s + \epsilon]$ and

[9] Adding a signal for the village is an interesting extension, which is left for future work.

a symmetric c.d.f. G_s. This more general approach will complicate the analysis since the prior and the posterior will no longer necessarily be from a conjugate family but would lead to the same results since the condition in (3.16) would, in fact, remain the same.

Also, instead of assuming that there is a sunspot state which is observed with an error, one can model strategic uncertainty by assuming that nature draws $(s_1, s_2) \sim H$, where s_i is the sunspot to bandit i and H is the joint cumulative distribution of the random vector (s_1, s_2). In the private sunspot case, the two sunspots s_1 and s_2 will not be perfectly correlated, that is, $s_i = s_{-i} + \epsilon_i$, where ϵ_i is small noise term. This alternative approach will lead to the same conclusions as those presented in the text.

Comparative statics. Observe that $q^{\mathrm{P}} = 1 - 4tc$, is decreasing in the cost that each bandit must pay in order to attack c and in the cost to produce output t. The effect of c on the probability of an attack makes intuitive sense: a higher c both increases the gain of successful attack and the cost of failing to attack the village jointly. At the same time, the model predicts that villages facing larger cots to produce the same amount of output would also face lower probability of an attack from the bandits. This mechanism operates through the player's expectations that any equilibrium with $q > 0$ must have $\theta(q) = 4c$ and therefore $q = 1 - 4tc$. I will revisit the comparative statics of the model in the next section where I relate the properties of the private sunspot framework with that of the global game.

3.4 Relation to Global Games

One popular approach in the literature for dealing with situations similar to the Bandit Game is to avoid sunspots altogether and instead re-formulate the Bandit Problem as a global game. The benefits of the global game approach are clear – it allows us to endogenize the probability of a coordinated attack. In fact, under certain conditions, the equilibrium of the global game version of the Bandit Game will be unique. There are costs as well, however, both of technical and conceptual nature. Specifically, the Bandit Game as it currently stands does not have upper dominance region. More importantly, however, the strategic choice of the village precludes the global game logic from applying unless we introduce some randomness. Nonetheless, as I show in this section, both of these issues can be remedied and the Bandit Game can be re-formulated as a global game application.

3.4.1 A modified Bandit Game

Previously the cost of attacking the village for each bandit was always assumed to be positive and fixed to $c > 0$. This implies that, regardless of the output chosen by the village, a bandit who attacks without the help of the other bandit would obtain zero output while still paying a positive cost of $c > 0$ and therefore attacking the village is never a strictly dominant action. In this section I modify the model by assuming that the attack cost c is a random variable, with cumulative distribution function F and domain in $[c_L, c_H]$, where $c_L < 0 < c_H$, and such that the expected cost to attack is equal to $\hat{c} > 0$. That is

$$\hat{c} \equiv \int_{c_L}^{c_H} c dF$$

The interpretation of the modified version of the model is the following. The village chooses to produce an output θ which, as before, the bandits will be able to obtain only if they attack jointly. In addition, nature draws the cost of attack according to the distribution function F and if the realized value is positive the village's defense will be strong and each bandit choosing to attack (regardless of the action of the other) suffers a positive cost of c. On the other hand, if the realized value for the cost of attack is negative then the village's defense is weak and each bandit who choose to attack will seize a positive amount of output equal to c separately from whether or not the bandits attack jointly and seize θ.

I will consider both the private sunspot version and the global game version of the bandit game. In the global game version of the model, the village is choosing θ before knowing the realization of the cost of attack. The village's choice of θ is observed perfectly by the bandits. Each bandit, in addition, receives a noisy signal of the cost to attack the village. Thus, in the global game version, the village's takes into account that the attack probability is a function of its choice of θ. In the private sunspot case, in contrast, the village is taking as given the strategies of the bandits, and therefore, the probability of an attack induced by them. The output produced by the village, moreover, is not observed by the bandits (unless they manage to seize it).

My goal in this section is to compare the probability of an attack on the village in the two versions of the model. I will focus primarily on the ex-post probability of an attack, that is, after the cost to attack c has been realized, since the two approaches would yield very different comparative statics in this case.

3.4.2 Private sunspots version

In this section, I consider the private sunspot version of the model. For given probability of an attack q the village chooses θ in order to maximize its expected payoff:

$$(1-q)\,\theta - t\frac{\theta^2}{2}$$

which yields a familiar first order condition, namely $\theta^P(q) = (1-q)/t$. We can consider two cases for the private sunspot equilibria depending on whether the bandits observe the specific realization of the cost of attack c before making their decisions.

If the bandits observe the realized value of c, then in addition to their private sunspot signal s_i, their strategies can be conditioned on the specific value of the cost to attack c. The set of private sunspot equilibria for given realization of the cost of attack c are presented below.

- If $c < 0$, then there is a unique equilibrium where $q^{\mathrm{P}} = 1$ and $\theta^{\mathrm{P}}\left(q^{\mathrm{P}}\right) = 0$.
- If $0 \leq c < \dfrac{1}{4t}$, then there are two equilibria – one where $q^{\mathrm{P}} = 0$ and $\theta^{\mathrm{P}}\left(q^{\mathrm{P}}\right) = 0$ and one where $q^{\mathrm{P}} = 1 - 4tc$ and $\theta^{\mathrm{P}}\left(q^{\mathrm{P}}\right) = 4c$.
- If $c > \dfrac{1}{4t}$, then there is a unique equilibrium where $q^{\mathrm{P}} = 0$ and and $\theta^{\mathrm{P}}\left(q^{\mathrm{P}}\right) = 0$.

Henceforth, I restrict attention to the private sunspot equilibria with a positive attack probability (if it exists). After the realization of c, the ex-post probability of an attack will be equal to 1 for $c < 0$, equal to $1 - 4tc$ for $0 \leq c < \dfrac{1}{4t}$ and equal to 0 for $c > \dfrac{1}{4t}$. This is shown on panel (a) in Figure 3.6.

On the other hand, if the cost of attack is unknown to the bandits at the time they must choose whether to attack the village, then the private sunspot equilibria would be characterized by Proposition 3.4, with the only difference being that we must substitute c with the expected cost of attack $\hat{c}$. That is, there are two private sunspot equilibria – one where the probability of an attack is zero $q^{\mathrm{P}} = 0$ and the village's output is $\theta^{\mathrm{P}} = 1/t$ and one where the probability of an attack is $q^{\mathrm{P}} = 1 - 4t\hat{c}$ and the village's output is $\theta^{\mathrm{P}} = 4\hat{c}$.

3.4.3 Global game version

In this section, I analyze the global game version of the Bandit Game. The village must choose the level of output θ before the cost of attack has been realized. For given choice of θ and for a given realization of c, the payoffs for the bandits are shown below.

	Attack	Not attack
Attack	$\theta/2 - c$, $\theta/2 - c$	$-c$, 0
Not attack	0, $-c$,	0, 0

Table 3.2: Payoff Matrix for given θ and c

The underlying complete information sub-game for fixed θ now has both upper and lower dominance regions. In particular, if $c < c_L(\theta) = 0$ then attacking will be the strictly dominant action for the bandits. On the other hand, if $c > c_H(\theta) = 2\theta$, then not attacking will be the strictly dominant action for the bandits.

Next, consider a perturbation of the original game where the bandits still observe the choice of θ perfectly, but the cost of an attack c is observed only with a small noise. Specifically, each bandit receives a private and independently distributed signal of the specific realization of c and decides whether to attack the village. For given realization of c, the signal for bandit i is the following:

$$c_i = c + n_i, \quad \text{where} \quad n_i \sim U\left[-n, n\right]$$

A strategy for bandit i is a choice of whether to attack $a_i = 1$ or not $a_i = 0$ for each realization of his private signal c_i, that is, $a_i(c_i) \in \{0\,1\}$. For fixed θ, the incomplete information version of the bandit game has a unique equilibrium, where each bandit attacks if and only if his signal c_i is below a threshold point $c^*(\theta)$, that is, $a_i(c_i) = 1$ iff $c_i < c^*(\theta)$ where $c^*(\theta)$ is the solution to:

$$\frac{1}{2}\left(\frac{\theta}{2} - c^*(\theta)\right) + \frac{1}{2}\left(-c^*(\theta)\right) = 0$$

In other words, for given θ, if $c_i = c^*(\theta)$ then the bandit is indifferent between choosing to attack and choosing not to attack when the other bandit will attack with probability equal to $1/2$. The above condition yields that the threshold point for given θ must equal $c^*(\theta) = \theta/4$.[10] In the limit as the noise in the bandits' signals approaches 0, i.e. $n \to 0$, the ex-ante probability that the bandits choose the same action goes to 1. Furthermore, as the noise in their private signals disappears, the

[10] Observe that the threshold point(s) determining the attack probability in the private sunspot version of the game are not related to the fundamentals of the environment. This is not the case in the global game version, where the strategies of the bandits have a threshold point $c^*(\theta)$ which is a direct function of the choice made by the village.

probability of an attack on the village approaches the probability that the cost of attacking is below $c^*(\theta)$. That is,

$$F\left(c^*(\theta)\right) = F\left(\frac{\theta}{4}\right)$$

Henceforth, in order to facilitate comparison with the private sunspot version, I restrict attention to the case where the noise in the bandits' signals is arbitrarily close to zero. The village will choose its level of output in order to maximize:

$$\left(1 - F\left(\frac{\theta}{4}\right)\right)\theta - t\frac{\theta^2}{2}$$

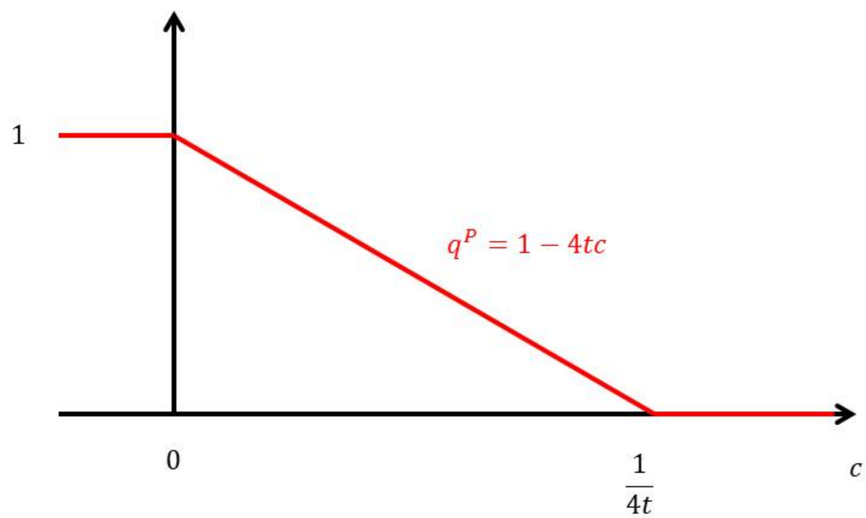
Differentiating the above expression with respect to θ and setting the derivative to zero, we obtain that the optimal choice of θ will be implicitly defined as the solution to the following equation:

$$\theta = \frac{1 - F\left(\frac{\theta}{4}\right)}{t + F'\left(\frac{\theta}{4}\right)\frac{1}{4}} \equiv g(\theta) \tag{3.17}$$

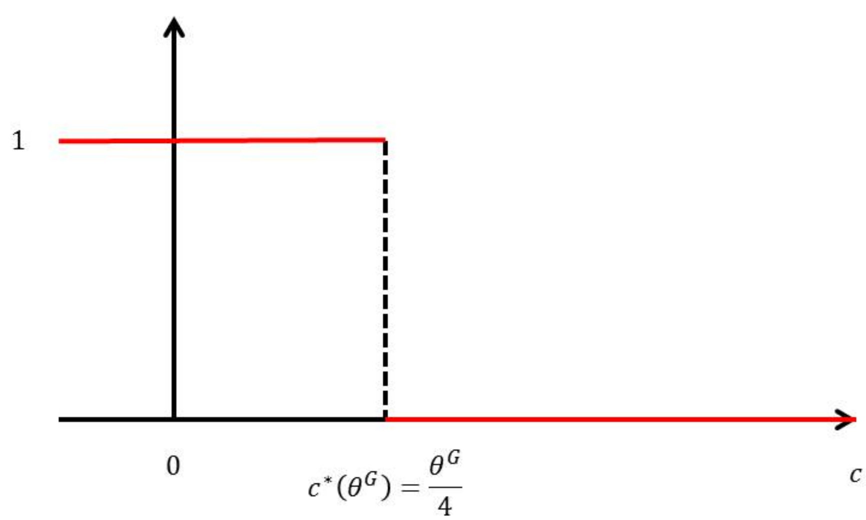
The equilibrium choice of θ is determined as the solution to $\theta^{\mathrm{G}} = g\left(\theta^{\mathrm{G}}\right)$ and the ex-ante equilibrium attack probably (i.e. before the realization of c) in the global game version of the Bandit Game is thus obtained from $q^{\mathrm{G}} = F\left(\frac{\theta^{\mathrm{G}}}{4}\right)$. On the other hand, the ex-post probability of an attack on the village (i.e. after the realization of c) will be equal to 1 for $c < c^*\left(\theta^{\mathrm{G}}\right) = \frac{\theta^{\mathrm{G}}}{4}$ and equal to 0 for $c > c^*\left(\theta^{\mathrm{G}}\right) = \frac{\theta^{\mathrm{G}}}{4}$. This is shown on panel (b) in Figure 3.6.

Comparative statics. The ex-post probability of an attack on the village is shown on panel (a) of Figure 3.6 for the private sunspot version of the model and on panel (b) for the global game version of the model. Observe that the two approaches to endogenizing the probability of an attack on the village yield markedly different comparative statics. In the private sunspot case, the ex-post attack probability is a linear and decreasing function of c (at least in the interval $[0,\ 1/4t]$). In the global game case, in contrast, the ex-post attack probability will be either 1 or 0. In other words, when the signals become arbitrarily precise, the global game select one of the two symmetric, pure-strategies Nash equilibria of the underlying model. When the cost c is below the cutoff $c^*\left(\theta^{\mathrm{G}}\right)$, the attack equilibrium is selected, and when the cost is above $c^*\left(\theta^{\mathrm{G}}\right)$ the no-attack equilibrium is selected. In contrast, the private-sunspots approach generates an equilibrium selection mechanism that assigns non-trivial probabilities to these two equilibria for a range of values of the cost c. In addition, the probability of an attack responds continuously to a change in the cost c (as shown in panel (a) of Figure 3.6), instead of jumping discontinuously as in panel (b). In this way, the private sunspots approach can be viewed as endogenously generating the type of equilibrium selection mechanism advocated by Ennis and Keister (2005a) for the analysis of government policy in models with complementarities and multiple equilibria.[11]

[11] In related work, Ennis and Keister (2005b) show how an equilibrium selection mechanism with these general properties can result from an adaptive learning process with boundedly-rational agents. The approach I take here, in contrast, is fully consistent with rational expectations.



(a) Ex-post probability of an attack on the village in the private sunspots version of the Bandit Game.



(b) Ex-post probability of an attack on the village in the global game version of the Bandit Game.

Figure 3.6: Ex-post probability of an attack on the village.

3.5 Generalizing the model

The Bandit Game is meant to capture games of regime change, where the benefit of "attacking" the regime is determined by the choice of a third player. Specifically, there are two players (speculators, depositors, bandits), each choosing between two actions (short sell the currency, run on the bank, attack the village). In addition, there is a third player (the government, the bank, the village) whose action is continuous (level of reserves, amount of liquidity, how much to produce). The choice of this third player directly affects the benefit and cost trade-off faced by the other two players. Finally, the environment is one where there is lack of commitment. This, in particular, entails that the aforementioned third player takes the strategies of the other two players as given when choosing his action. In this section, I show how the logic of the example from the previous section can readily be extended to more general payoff functions.

3.5.1 Payoffs

For simplicity, I continue to refer to the players facing the binary choice as the bandits and to the player choosing θ as the village.[12] In addition, I continue to use the same setup for the sunspot state, namely $s \sim U[0,1]$ and each bandit receives a signal s_i which, for given s, is distributed as $U[s-\epsilon,\, s+\epsilon]$. So, suppose we have 3 players: bandits i, bandit j and the village. The actions of the three players are summarized in a 3-tuple:

$$(a_i,\, a_j,\, \theta) \in \{0,\, 1\} \times \{0,\, 1\} \times \Theta$$

where $a_i \in \{0,\, 1\}$ is the choice of bandit i ($a_i = 1$ corresponds to attacking the village and $a_i = 0$ corresponds to not attacking) and $\theta \in \Theta = \left[0,\, \bar{\theta}\right]$ is the village's choice of output level. All three players made their choices simultaneously and independently

[12]Moreover, in order to simplify the exposition, I will restrict attention to two bandits. The extension to more than two bandits is left for future work.

of each other. The bandits' payoff are symmetric and given by

$$u : \{0, 1\} \times \{0, 1\} \times \Theta \to \mathbb{R} \tag{3.18}$$

where $u(a_i, a_j, \theta)$ is bandit i's payoff from choosing a_i given that the choice of the other bandit is a_j and the village's choice is θ. The payoff from choosing not to attack is normalized to zero

$$u(0, a_j, \theta) = 0 \quad \text{for} \quad \text{all} \quad a_j \text{ and } \theta \tag{3.19}$$

I impose the following assumptions.

(A1): $u(1, 1, \theta) - u(1, 0, \theta) > 0$ i.e. there is a benefit of coordination.

(A2): $u(1, 1, \theta) - u(1, 0, \theta)$ is strictly increasing in θ.

(A3): $u(a_i, a_j, \theta)$ is continuous function of θ.

The payoff for the village is given by

$$W : \{0, 1\} \times \{0, 1\} \times \Theta \to \mathbb{R} \tag{3.20}$$

where the payoff for the village $W(a_1, a_2, \theta)$ depends on the choice of the bandits and the choice of the village.

3.5.2 Strategies

A strategy for bandit i is a mapping from his private signal to a decision of whether to attack the village or not, that is $a_i(s_i) \in \{0, 1\}$. A strategy for the village is a choice of θ. Suppose that each bandit attacks when his signal is in a given set $A \subseteq [0, 1]$. That is, $a_i(s_i) = 1$ if and only of $s_i \in A$. From the perspective of the village, the probability distribution over the actions of the bandits will be

$$q_{11} = \Pr\left[s_i \in A,\, s_j \in A\right], \quad q_{10} = \Pr\left[s_i \in A,\, s_j \notin A\right], \quad q_{01} = \Pr\left[s_i \in A,\, s_j \notin A\right]$$

and $q_{00} = 1 - q_{11} - q_{01} - q_{10}$. Where q_{11} is the probability both bandits choose to attack, q_{00} is the probability that both bandits choose not to attack and $q_{10} + q_{01}$ is the probability that only one of them chooses to attack. The expected payoff for the village becomes

$$q_{11}W(1,\, 1,\, \theta) + q_{10}W(1,\, 0,\, \theta) + q_{01}W(0,\, 1,\, \theta) + q_{00}W(0,\, 0,\, \theta)$$

As the noise in the private signal converges to zero, we have $q_{10} = q_{10} \to 0$. Then by denoting $q \equiv q_{11}$, the probability that both bandits attack is simply the probability that the sunspot state is in the set A i.e. $q = \Pr\left[s \in A\right]$. For given q, the objective of the village is to choose θ in order to maximize its expected payoff $\hat{W}\left(q,\, \theta\right)$. Let $\hat{\theta}(q)$ denote the solution to this problem. That is,

$$\hat{\theta}(q) = \underset{\theta \in \Theta}{argmax} \quad \{(1-q)W(1,\, 1,\, \theta) + qW(0,\, 0,\, \theta)\} \tag{3.21}$$

I impose the following assumptions on the optimal choice of output $\hat{\theta}(q)$:

(A4): For each $q \in [0, 1]$ the solution to (3.21) exist and is unique.

(A5): The level of output set by the village is inversely related to the probability of an attack, $\dfrac{\partial \hat{\theta}(q)}{\partial q} < 0$.

According to **(A5)**, The game has a zero sum property in the sense that output obtained by bandits is an output lost by the village and the later wants to avoid being attacked by the bandits. Next, I turn to the bandits. Define the expected net gain of choosing to attack, $a_i = 1$, given that the other bandit chooses $a_{-i} = 1$ with probability β:

$$f(\beta,\, \theta) \equiv \beta u(1,\, 1,\, \theta) + (1-\beta)u(1,\, 0,\, \theta) \tag{3.22}$$

Given the the strategies of the bandits, we have $\beta(s_i) = \Pr\left[s_j \in A\,|s_i\right]$. From **(A1)** - **(A3)** it follows that the function $f(\beta,\, \theta)$ is continuous and increasing in β and in θ.

$$\frac{\partial f(\beta,\, \theta)}{\partial \beta} > 0, \quad \frac{\partial f(\beta,\, \theta)}{\partial \theta} > 0$$

Define θ^* and q^* to be the solution to the following system of equations.

$$f\left(\frac{1}{2},\, \theta^*\right) = \frac{u\left(1,\, 1,\, \theta^*\right) + u\left(1,\, 0,\, \theta^*\right)}{2} = 0 \tag{3.23}$$

$$\theta^* = \underset{\theta\in\Theta}{argmax} \quad \left\{(1-q^*)\, W\left(1,\, 1,\, \theta^*\right) + q^* W\left(0,\, 0,\, \theta^*\right)\right\} \tag{3.24}$$

That is, θ^* is the value of the village's output which makes each bandit indifferent between choosing to attack and choosing not to attack if the other bandit is mixing with equal probability. Given the properties of the function f, there is at most one solution to the equation $f\left(\frac{1}{2},\, x\right) = 0$. At the same time, if the village anticipates an attack with probability q^* then it becomes optimal to set output to $\hat{\theta}\left(q^*\right) = \theta^*$.

3.5.3 Equilibrium

Now we are ready to characterize the private sunspot equilibria of this generalized version of the bandits' game.

Proposition 3.5. *Equilibria of the private sunspot game*

(i) If $f\left(0,\, \hat{\theta}(0)\right) \leq 0$ then there exist an equilibrium where the probability of an attack is $q = 0$.

(ii) If $f\left(1,\, \hat{\theta}(1)\right) \geq 0$ then there exist an equilibrium where the probability of an attack is $q = 1$.

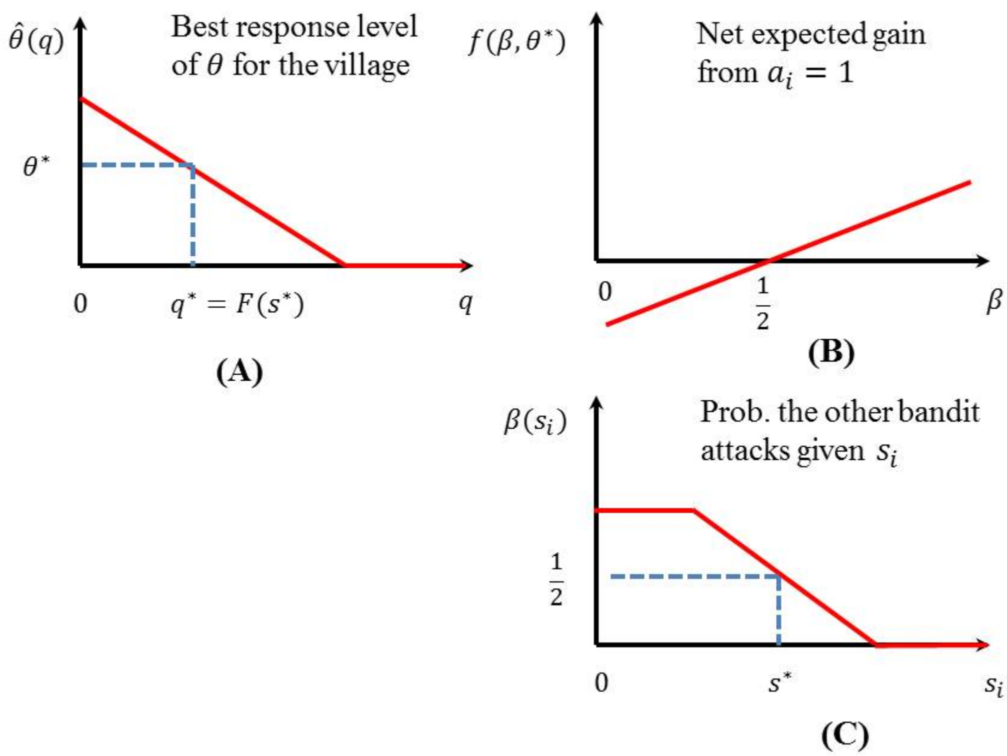


Figure 3.7: Private Sunspot Equilibrium with positive attack probability.

(iii) If θ^* and q^* are the solution to (3.23) *and* (3.24)*, then there exist an equilibrium where the probability of an attack is q^* and the output produced by the village is θ^*.*

Proof. Consider a profile of strategies for the bandits $a_i(s_i) = 1$ iff $s_i \in A$ for $i = 1, 2$. This profile induces a probability of an attack equal to q.

Suppose that the bandits never attack i.e. $a_1(s_i) = 0$ for all s_i. The output produced by the village will be equal to $\hat{\theta}(0)$. This is an equilibrium if $f\left(0,\, \hat{\theta}(0)\right) \leq 0$. If the previous inequality were not true, that is if $f\left(0,\, \hat{\theta}(0)\right) > 0$, then each bandit want to deviate and choose $a_i(s_i) = 1$ and therefore the no attack equilibrium will not exist.

Suppose that both bandits always attack i.e. $a_i(s_i) = 1$ for all s_i. The village will then set its output to $\hat{\theta}(1)$. The certain attack can be sustained in equilibrium only if $f\left(1,\, \hat{\theta}(1)\right) \geq 0$. Otherwise, if $f\left(1,\, \hat{\theta}(1)\right) < 0$ then each bandit wants to deviate and choose $a_i(s_i) = 0$ and therefore the certain attack equilibrium will not exist.

Suppose that the attack probability q is in the interval $(0, 1)$. Then set A must contain at least one threshold point and if the signal for bandit i equals a threshold point then he expects the other to attack with probability $\frac{1}{2}$ and must be indifferent between $a_i = 1$ and $a_i = 0$. That is,

$$f\left(\frac{1}{2},\, \hat{\theta}(q)\right) = 0$$

Given **(A1) - (A4)** this condition will be satisfied only if $q = q^*$, where q^* and $\hat{\theta}(q^*)$ are characterized by (3.23) and (3.24). Moreover, since $f\left(\beta,\, \theta^*\right) > f\left(\frac{1}{2},\, \theta^*\right) = 0$ for $\beta > \frac{1}{2}$ and $f\left(\beta,\, \theta^*\right) < f\left(\frac{1}{2},\, \theta^*\right) = 0$ for $\beta < \frac{1}{2}$, it follows that the bandits will be best responding by switching their action around any of the threshold points of the set A. $\square$

The determination of the equilibrium where the output produced is θ^* and the probability of an attack is q^* is illustrated on Figure (3.7). Specifically, assume that each bandit follows a threshold strategy and chooses to attack the village only if his private signal s_i is less or equal to s^*. Panel A depicts the best response of the village, which based on assumptions **(A4)** - **(A5)**, is decreasing towards zero as a function of q. As the noise in the private signals approaches zero, ex ante the village anticipates that either both bandits will attack (an event which occurs with probability q^*) or there will be no attack (an event which occurs with probability $1 - q^*$) and best responds with an output of θ^* as shown on panel A on the figure. Panel B depicts the net expected gain from choosing to attack given that the other bandit attacks with probability β and the village's output is expected to be θ^*. Assumptions **(A1)** – **(A3)** imply that the function $f(\beta, \theta^*)$ is increasing and crosses zero at most once, which when $\theta = \theta^*$will happen exactly at $\frac{1}{2}$. Finally, panel C shows the posterior probability assign by bandit i to the event that the other bandit chooses to attack.

3.6 Conclusion

This chapter proposed a new approach to endogenizing the probability of a self-fulfilling event. In particular, the idea is to introduce a small amount of strategic uncertainty into the original game of coordination. The way this was done, however, is not through an imperfectly observed fundamental state, but rather through imperfectly observed sunspot state. The private sunspot approach was illustrated within the context of a specific example where two bandits want to coordinate on attacking a village in order to seize whatever has been produced there. The village, in turn, forms expectations with respect to the probability of an attack and chooses its output accordingly. The attack on the village is entirely self-fulfilling and yet the private sunspots approach allowed us to derive the probability of this attack as a function of the parameters of the model. The reason is that the private sunspot approach introduces strategic uncertainty, which would then require the strategies of the bandits to satisfy an additional equilibrium condition. This condition is instrumental in pinning down the equilibrium probability of an attack on the village.

The private sunspot approach seems especially suited for games where there is lack of pre-commitment. Thus, in our example, the village would decide how much to produce ex-post, taking the probability of experiencing an attack as given. Finally, though, the private sunspot approach was illustrated with a specific example, the logic applies more generally. In particular, strategic interactions where there is a benefit of coordination and this benefit depends on the choice made by a third player can easily be mapped into the private sunspot framework. Bank runs, currency attacks and government defaults would appear to be natural candidates for such an exercise. Furthermore, extending the model to a continuum of bandits seems an obvious next step.

Bibliography

[1] Acemoglu, D. (2011) Thoughts on inequality and the financial crisis. In Presentation held at the American Economic Association.

[2] Angeletos, G (2008), "Private Sunspots and Idiosyncratic Investor Sentiment," NBER Working Papers 14015, National Bureau of Economic Research, Inc.

[3] Andolfatto, D., Nosal, E. and Wallace, N. (2007) "The role of independence in the Green-Lin Diamond-Dybvig model," Journal of Economic Theory 137: 709-715.

[4] Andolfatto, D., Nosal, E., and Sultanum, B. (2016) "Preventing bank runs," mimeo., April.

[5] Allen, F. and Gale, D. (2004) "Financial fragility, liquidity, and asset prices" *Journal of the European Economic Association,* **2**, 1015-1048.

[6] Azariadis, Costas (1981), "Self-fulfilling Prophecies," Journal of Economic Theory 25 380-96.

[7] Argenziano, R. (2008). "Differentiated Networks: Equilibrium and Efficiency." RAND Journal of Economics 39:747-769.

[8] Basel Committee on Banking Supervision (BCBS, 2013) "Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools" Available at www.bis.org/publ/bcbs238.pdf

[9] Benhabib, J., and Farmer R., (1994), "Indeterminacy and Increasing Returns," Journal of Economic Theory 63, 19-41.

[10] Bianchi, J. (2012), "Efficient bailouts?" NBER Working Paper No. 18587.

[11] Boyd, J.H., Chang, C. and Smith, B.D. (2002) "Deposit insurance: A reconsideration" *Journal of Monetary Economics,* **49,** 1235-1260.

[12] Brown, C. and Dinç, S, (2005), The Politics of Bank Failures: Evidence from Emerging Markets, *The Quarterly Journal of Economics*, 120, issue 4, p. 1413-1444

[13] Burdzy, Krzysztof, and David M. Frankel. (2005). "Shocks and Business Cycles." Advances in Theoretical Economics vol. 5, iss. 1, paper no. 2.

[14] Calomiris, C.W., and Kahn, C.M. (1991) "The Role of Demandable Debt in Structuring Optimal Banking Arrangements" *American Economic Review,* **81**, 497-513.

[15] Carlsson, H. and E. van Damme (1993), “Global Games and Equilibrium Selection,” Econometrica, 61, 989–1018.

[16] Calvo, G. (1988), “Servicing the public debt: The role of expectations,” The American Economic Review, 647-661

[17] Chari, V.V. and Kehoe, P.J. (2015) “Bailouts, time inconsistency, and optimal regulation: A macroeconomic view” Federal Reserve Bank of Minneapolis Staff Report 481, January.

[18] Cooper, R. and Kempf, H. (2013) “Deposit insurance and bank liquidation without commitment: Can we sleep well?” NBER Working Paper No. 19132.

[19] Cooper, Russell (1999), Coordination Games: Complementarities and Macroeconomics, Cambridge, UK: Cambridge University Press.

[20] Cooper, R. and Ross, T.W. (1998) “Bank runs: liquidity costs and investment distortions” *Journal of Monetary Economics,* **41**, 27-38.

[21] Chamley, C., (1999). ”Coordinating Regime Switches.” Quarterly Journal of Economics 114:869-905.

[22] Diamond, D.W. and Dybvig, P.H. (1983) “Bank runs, deposit insurance, and liquidity” *Journal of Political Economy,* **91**, 401-419.

[23] Donges, J. and F. Heinemann (2000), “Competition for Order Flow as a Coordination Game,” Center for Financial Studies, Frankfurt, Germany.

[24] EFTA Court (2013), ”Judgement of the Court (on deposit-guarantee schemes, Obligation of result, Emanation of the State, Discrimination)”

[25] Ennis, H.M. and Keister, T. (2005a) “Government Policy and the Probability of Coordination Failures” European Economic Review, Vol. 49 (2005) 939-973.

[26] Ennis, H.M. and Keister, T. (2005b) “Optimal fiscal policy under multiple equilibria” Journal of Monetary Economics, 52, 1359–1377

[27] Ennis, H.M. and Keister, T. (2009a) “Bank runs and institutions: The perils of intervention” *American Economic Review,* **99**, 1588-1607.

[28] Ennis, H.M. and Keister, T. (2009b) “Run equilibria in the Green-Lin model of financial intermediation” *Journal of Economic Theory,* **144**, 1996-2020.

[29] Ennis, H.M. and Keister, T. (2010) “Banking panics and policy responses” *Journal of Monetary Economics,* **57**, 404-419.

[30] Farhi, E. and Tirole, J. (2012) “Collective moral hazard, maturity mismatch and systemic bailouts” *American Economic Review,* **102**, 60-93.

[31] Fitoussi, J., Saraceno, R. (2010) Europe: How Deep Is a Crisis? Policy Responses and Structural Factors Behind Diverging Performances, Journal of Globalization and Development, vol.1, issue1, article 17

[32] Freeman, Scott (1988) "Banking as the provision of liquidity," Journal of Business 61: 45-64.

[33] Gale, D. and Vives, X. (2002) "Dollarization, bailouts, and the stability of the banking system" *Quarterly Journal of Economics,* **117**, 467-502.

[34] Goldstein, I. and Pauzner, A. (2005) "Demand deposit contracts and the probability of bank runs" *Journal of Finance,* **60**, 1293-1327.

[35] Green, E.J. (2010) "Bailouts" Federal Reserve Bank of Richmond *Economic Quarterly,* **96**,11-32.

[36] Green, E.J. and Lin, P. (2000) "Diamond and Dybvig's classic theory of financial intermediation: What's missing?" Minneapolis Fed *Quarterly Review,* **24** (Winter), 3-13.

[37] Green, E.J. and Lin, P. (2003) "Implementing efficient allocations in a model of financial intermediation" *Journal of Economic Theory,* **109**, 1-23.

[38] Gu, C. (2011) "Nosy Sunspots and Bank Runs", Macroeconomic Dynamics, 15, 2011, 398–418.

[39] Harsanyi, J., (1974). "An equilibrium-point interpretation of stable sets and a proposed alternative definition", Management Science, 20, 1472–1495.

[40] Keister T. (2016) "Bailouts and Financial Fragility" *Review of Economic Studies*, **83,** 704–736.

[41] Morris, S. and H. Shin (2002) "Global Games: Theory and Applications," Advances in Economic Theory, Proceedings of the Econometric Society Ninth World Congress.

[42] Kumhof, M., Rancière, R. (2011). "Inequality, leverage and crises" IMF working Paper 10/268.

[43] Kroszner, R. and Strahan, P. (1996). "Regulatory Incentives and the Thrift Crisis: Dividends, Mutual- to-Stock Conversions, and Financial Distress." Journal of Finance **51**, 1285– 320.

[44] Martin, A. (2006) "Liquidity provision vs. deposit insurance: preventing bank panics without moral hazard" *Economic Theory,* **28**, 197–211.

[45] Nosal, Jaromir and Guillermo Ordoñez (2013) "Uncertainty as commitment," NBER Working Paper No. 18766.

[46] Peck, J. and Shell, K. (2003) "Equilibrium bank runs" *Journal of Political Economy,* **111**, 103-123.

[47] Fukao, K. (1994), "Coordination Failures under Incomplete Information and Global Games," Discussion Paper Series A 299, The Institute of Economic Research, Hitotsubashi University, Kunitachi, Tokyo, Japan.

[48] Harsanyi, J. and R. Selten (1988). AGener al Theory of Equilibrium Selection in Games, Cambridge, MA: MIT Press.

[49] Postlewaite, A. and Vives, X. (1987) "Bank runs as an equilibrium phenomenon" *Journal of Political Economy,* **95**, 485-491.

[50] Kajii, A. and S. Morris (1997a), "The Robustness of Equilibria to Incomplete Information," Econometrica, 65, 1283–1309.

[51] Obstfeld, M. (1996), "Models of Currency Crises with Self-Fulfilling Features," European Economic Review, 40, 1037–1047.

[52] Oyama, Daisuke. (2004). "Booms and Slumps in a Game of Sequential Investment with the Changing Fundamentals." Japanese Economic Review 55:311-20.

[53] Toxvaerd, Flavio. (2008). "Strategic Merger Waves: A Theory of Musical Chairs." Journal of Economic Theory 140:1-26.

[54] Ray, D., Vohra, R., (2015). "Coalition formation," Handbook of Game Theory with Economic Applications, in: Handbook of Game Theory with Economic Applications, volume 4, chapter 5, pages 239-326 Elsevier.

[55] Ranciere, Romain and Aaron Tornell (2011) "Financial liberalization: Efficiency gains and black holes," mimeo., October.

[56] Rogoff, K. and Sibert, A (1988) "Elections and Macroeconomic Policy Cycles," *Review of Economic Studies* **55**, Jan. 1988, 1-16

[57] Rogoff, Kenneth "The Optimal degree of commitment to an intermediate monetary target," Quarterly Journal of Economics 100 (November 1985), 1169-1189.

[58] Rubinstein, A. (1989), "The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge," American Economic Review, 79, 385–391.

[59] Stiglitz, J.E. (2012) The Price of Inequality: How Today's Divided Society Endangers Our Future. W. W. Princeton NJ: Norton & Company.

[60] Stiglitz, J. (2009) "The global crisis, social protection and jobs," International Labour Review, Vol. 148, No. 1–2, pp.1-13

[61] Sibert, Ann (2012) "Deposit insurance after Iceland and Cyprus," VoxEU.org, 2 Apr 2012.

[62] Smith, B.D. (1988). Legal restrictions, 'sunspots', and Peel's bank act: the real bills doctrine versus the quantity theory reconsidered. Journal of Political Economy 96: 3-19.

[63] Shell, K. (2008) "Sunspot Equilibrium", The New Palgrave: A Dictionary of Economics, 2nd Edition (L. Blume and S. Durlauf, eds.), New York: Palgrave Macmillan, 2008

[64] Wallace, N. (1988) "Another attempt to explain an illiquid banking system: the Diamond and Dybvig model with sequential service taken seriously" Minneapolis Fed *Quarterly Review,* **12** (Fall), 3-16.

[65] Wallace, N. (1990) "A banking model in which partial suspension is best" Minneapolis Fed *Quarterly Review,* **14** (Fall), 11-23.

[66] Woodford, M. (1990). Learning to believe in sunspots. Econometrica 58: 227-307.

[67] M. Woodford, (1986) Stationary sunspot equilibria in a finance constrained economy, J. Econ. Theory 40, 128 137.